

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-520744

(P2004-520744A)

(43) 公表日 平成16年7月8日(2004.7.8)

(51) Int.Cl.⁷

H04N 7/32

F I

H04N 7/137

Z

テーマコード (参考)

審査請求 未請求 予備審査請求 有 (全 62 頁)

(21) 出願番号 特願2002-560391 (P2002-560391)
(86) (22) 出願日 平成14年1月25日 (2002.1.25)
(85) 翻訳文提出日 平成15年7月28日 (2003.7.28)
(86) 国際出願番号 PCT/FR2002/000322
(87) 国際公開番号 W02002/060184
(87) 国際公開日 平成14年8月1日 (2002.8.1)
(31) 優先権主張番号 01/01174
(32) 優先日 平成13年1月26日 (2001.1.26)
(33) 優先権主張国 フランス (FR)

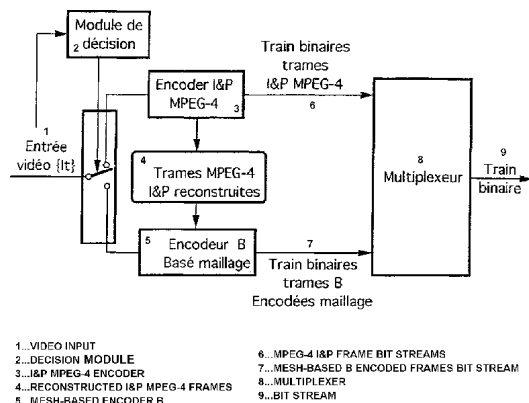
(71) 出願人 591034154
フランス テレコム
FRANCE TELECOM
フランス国、75015 パリ、プラス・
ダルレ、6
(74) 代理人 100099623
弁理士 奥山 尚一
(74) 代理人 100096769
弁理士 有原 幸一
(74) 代理人 100107319
弁理士 松島 鉄男
(72) 発明者 サンソン, アンリ
フランス国、35690 アシーニュ, ア
レ・デ・ロソー 12

最終頁に続く

(54) 【発明の名称】 画像の符号化と復号のための方法及び装置、それに関連する用途

(57) 【要約】

画像を符号化するための方法であって、それぞれ異なる最適化基準に基づいてビデオシーケンスの少なくとも1つの画像のコンプレッションを最適化する少なくとも2つの方式の画像符号化を選択的に実行する。



【特許請求の範囲】

【請求項 1】

画像を符号化するための方法において、それぞれ異なる最適化基準に基づいてビデオシーケンスの少なくとも 1 つの画像のコンプレッションを最適化する少なくとも 2 つの方式の画像符号化を選択的に実行する、ことを特徴とする画像符号化方法。

【請求項 2】

前記符号化方式の中の 1 つの方式の選択を表す情報は復号器によって、

- ・符号化と復号の段階において知られる、予め定められた選択と、
- ・少なくとも一部の符号化された画像データを含むデータフローに含まれる前記選択の表示情報と、
- ・符号化された画像データとは無関係なデータフローに含まれる前記選択の表示情報と、
- ・前記復号器が固有の仕方で前記選択を判定すること、から成るグループに属する技術の中の少なくとも 1 つに基づいて知られる、ことを特徴とする請求項 1 に記載の画像符号化方法。

10

【請求項 3】

前記画像に適用する符号化方式を、少なくとも、

- ・画像の光度表現を十分に最適化する第 1 の符号化と、
- ・少なくとも 2 つの画像にわたる動きの表現を十分に最適化する第 2 の符号化と、の中から選択するステップを具備する、ことを特徴とする請求項 1 及び 2 のいずれかに記載の画像符号化方法。

20

【請求項 4】

前記第 2 の符号化は、前記第 1 の符号化を使用して符号化された少なくとも 1 つの以前の画像及び／または少なくとも 1 つの以後の画像を考慮に入れる、ことを特徴とする請求項 3 に記載の画像符号化方法。

【請求項 5】

前記第 2 の符号化は、前記第 1 の符号化を使用して符号化されたすぐ前の画像から計算される動きベクトル・フィールド及び／または前記第 1 の符号化を使用して符号化されたすぐ後の画像から計算される動きベクトル・フィールドを考慮に入れる、ことを特徴とする請求項 4 に記載の画像符号化方法。

【請求項 6】

前記動きベクトル・フィールドはメッシュ網に適用される、ことを特徴とする請求項 5 に記載の画像符号化方法。

30

【請求項 7】

前記動きベクトル・フィールドは、前記第 2 の符号化を使用して符号化された画像に関係した、推論された動きベクトル・フィールドを決定するために使用される、ことを特徴とする請求項 5、6 のいずれかに記載の画像符号化方法。

【請求項 8】

前記第 2 の符号化は、画像 I_{t1} と次の画像 I_{t2} の間の事前動き推定と、該動き推定の間に得られたかつ対称変換を引き起こさない動きベクトルが、 m が $[t1, t2]$ に属するとしたときに少なくとも 1 つの画像 I_{m1} を補間するために $0 < k < 1$ のスカラー値 $k = m / (t1 + t2)$ によって重み付けされる動き補償の段階とを実行する、ことを特徴とする請求項 5 乃至 7 のいずれかに記載の画像符号化方法。

40

【請求項 9】

前記第 2 の符号化は、画像 I_{t2} と前の画像 I_{t1} の間の事後動き推定と、該動き推定の間に得られたかつ対称変換を引き起こさない動きベクトルが、 m が $[t1, t2]$ に属するとしたときに少なくとも 1 つの画像 I_{m2} を補間するために $0 \leq k' \leq 1$ のスカラー値 $k' = 1 - m / (t1 + t2)$ によって重み付けされる動き補償の段階とを実行する、ことを特徴とする請求項 5 乃至 7 のいずれかに記載の画像符号化方法。

【請求項 10】

$0 \leq a \leq 1$ として $I_m = a E1 + (1 - a) E2$ なる推定画像 I_m を得るために事後推定と

50

事前推定を実行する、ことを特徴とする請求項 8、9 のいずれかに記載の画像符号化方法。

【請求項 11】

前記選択段階は、N に 1 つの画像が前記第 1 の符号化を使用して符号化される固定因子 N のサブサンプリングの実施に基づく、ことを特徴とする請求項 3 乃至 10 のいずれかに記載の画像符号化方法。

【請求項 12】

N は 2 より大きい、ことを特徴とする請求項 11 に記載の画像符号化方法。

【請求項 13】

N は可変的である、ことを特徴とする請求項 11、12 のいずれかに記載の画像符号化方法。 10

【請求項 14】

前記第 1 の符号化は、画像ブロック上の変換とブロックによる時間的予測を実行する、ことを特徴とする請求項 3 乃至 11 のいずれかに記載の画像符号化方法。

【請求項 15】

前記第 1 の符号化は M P E G - 4 符号化または H 2 6 符号化である、ことを特徴とする請求項 14 に記載の画像符号化方法。

【請求項 16】

M P E G - 4 符号化または H 2 6 符号化によって得られる画像はタイプ I (intra) 及び／またはタイプ P (predictable) の画像を含む、ことを特徴とする請求項 15 に記載の画像符号化方法。 20

【請求項 17】

前記第 2 の符号化は、M を 1 以上として M 階の階層型メッシュ網に基づく、ことを特徴とする請求項 3 乃至 16 のいずれかに記載の画像符号化方法。

【請求項 18】

前記メッシュ網は三角形型である、ことを特徴とする請求項 17 に記載の画像符号化方法。

【請求項 19】

被覆領域の管理段階を具備する、ことを特徴とする請求項 17、18 のいずれかに記載の画像符号化方法。 30

【請求項 20】

別個の伝送経路で伝送することができる少なくとも 2 つのデータフローを作り出す、ことを特徴とする請求項 1 乃至 19 のいずれかに記載の画像符号化方法。

【請求項 21】

前記データフローは、
グローバルヘッダと、
前記第 1 の符号化に基づいて符号化された画像データと、
前記第 2 の符号化に基づいて符号化された画像データと、
を含むグループに属する、ことを特徴とする請求項 20 に記載の画像符号化方法。

【請求項 22】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された画像信号を復号するための画像復号方法。 40

【請求項 23】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された画像信号を符号化するための画像符号化装置。

【請求項 24】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された画像信号を復号するための画像復号装置。

【請求項 25】

符号化において実行されたのと同様に、ベクトル・フィールドの少なくとも一部及び／ま 50

たは被覆領域の少なくとも一部を判定するための手段を具備する、ことを特徴とする請求項 24 に記載の画像復号装置。

【請求項 26】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された少なくとも 1 つの画像信号を記憶するための記憶装置。

【請求項 27】

請求項 1 乃至 17 のいずれかに記載された符号化方法を使用して符号化された画像信号を符号化、伝送及び／または復号するためのシステム。

【請求項 28】

前記符号化方式の中の 1 つの方式の選択を表す情報は復号器によって、
・符号化と復号の段階において知られる、予め定められた選択と、
・少なくとも一部の符号化された画像データを含むデータフローに含まれる前記選択の表示情報と、
・符号化された画像データとは無関係なデータフローに含まれる前記選択の表示情報と、
・前記復号器が固有の仕方で前記選択を判定することと、
から成るグループに属する技術の中の少なくとも 1 つに基づいて知られる、ことを特徴とする請求項 27 に記載のシステム。

【請求項 29】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された画像信号を符号化及び／または復号するためのコンピュータ・プログラム製品。

【請求項 30】

請求項 1 乃至 21 のいずれかに記載された符号化方法を使用して符号化された画像信号を符号化及び／または復号するためのコンピュータ・プログラムを格納したデータ記憶媒体。

【請求項 31】

請求項 1 乃至 21 のいずれかに記載された符号化方法に基づいて符号化されたデータを含む画像データ信号。

【請求項 32】

請求項 1 乃至 21 のいずれかに記載された符号化方法が働いたか否かを特定する少なくとも 1 つの指標を含む請求項 31 に記載の画像データ信号。

【請求項 33】

ビデオシーケンスの先頭及び／またはそれぞれの信号フレーム内に、フレーム構造を特定するデータ要素を含む、ことを特徴とする請求項 31、32 のいずれかに記載の画像データ信号。

【請求項 34】

前記第 2 の符号化を使用して符号化されたシーケンスは、この第 2 の符号化に基づいて符号化されたフレームの数を特定するヘッダから始まる、ことを特徴とする請求項 31 乃至 33 のいずれかに記載の画像データ信号。

【請求項 35】

別個の伝送経路で伝送することができる少なくとも 2 つのデータフローを含む、ことを特徴とする請求項 31 乃至 34 のいずれかに記載の画像データ信号。

【請求項 36】

前記データフローは、
グローバルヘッダと、
前記第 1 の符号化に基づいて符号化された画像データと、
前記第 2 の符号化に基づいて符号化された画像データと、
を含むグループに属する、ことを特徴とする請求項 35 に記載の画像データ信号。

【請求項 37】

請求項 1 乃至 21 のいずれかに記載された符号化方法を、
デジタル TV と、

I Pネットワーク上のリアルタイム・ビデオと、
移動体へのネットワーク上のリアルタイム・ビデオと、
画像データ保存と、を含むグループに属する領域の中の少なくとも1つの領域に適用することの特徴とする用途。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、動画像シーケンス、特にビデオのビットレートを減少させることによる符号化の技術に関する。

【背景技術】

【0002】

多数のビデオ符号化の用途が存在する。その中には（限定的なリストではないが）、

- ・デジタルTV伝送、
- ・I P、移動体通信といった数タイプのネットワークにわたるリアルタイムビデオ伝送（“ストリーミングI P”）、
- ・ビデオのコンピュータストレージ、等が含まれる。

【0003】

本発明は、特にMPEG型符号化を実行するシステムの枠組み内において適用可能である。MPEG型符号化とは、固定サイズが多いが可変サイズもあり得るハードブロック構造に基づく時間的予測（temporal prediction）と離散コサイン変換（discrete cosine transformation）に基づく符号化を意味する。この符号化ファミリーの2つの代表的な標準は、バージョン1からバージョン4までのMPEG-4標準とバージョン2までのITU-T/H. 263である。本発明は、CCITT H26リコメンデーション（例えば対応するVCEG-N83 d1ドキュメント参照）の枠組みにおいても適用可能である。

【0004】

現在提案されているビデオの符号化と復号のスキームは、以下の2つのカテゴリに分けられる。1つは、

- ・ISO/MPEGまたはITU-Tのいずれかによって標準化された符号化で、すべて同じタイプの技術（ブロック構造に基づく時間的予測と離散コサイン変換）に基づくものと、もう1つは、

- ・ウェーブレット符号化、地域的符号化（regional encoding）、フラクタル符号化、メッシュ符号化（meshing encoding）等の大きな一団の技術を必要とする研究所によって提案された開発中の符号化、である。

【0005】

現在、MPEG-4符号化は、標準化された符号化という点のみならず、公開された符号化という点で、業者間標準（trade standard）と考えられている。

【0006】

MPEG-4型またはITU-T/H. 263++型の符号化は、特に全ての符号化計算及び操作のための支援として使用される固定サイズのハードブロック構造が原因で限界に達していると考えられている。特に、シーケンス内の画像の時間的予測は十分に活用されていない。

【0007】

さらに、公開されている代わりの符号化は十分な度合いの最適化をまだ達成していない。

【0008】

このため、通常において符号器は、符号化されたビデオシーケンスを低いビットレートで得られるようにするために、画像のサイズを小さくし、オリジナルのビデオシーケンスの時間的なサブサンプルを生成する。それでもなお、この第2の技術はサブサンプルのレベルに応じたユーザにとっては多かれ少なかれ問題のある不規則な動き（erratic movements）を修復しなければならない不便さがある。

【0009】

10

20

30

40

50

これらの不規則な動きを避けるためには、復号器において時間的補間 (temporal interpolation) を使用して欠けた (符号化されていない) 画像が再生されなければならない。

【0010】

それでもなお、画像の時間的補間に関する現在の技術では、とりわけそれらの技術が単一の復号器において実施されるときには、満足な結果は得られない。実際、これらの技術は、1ブロックの全てのピクセルに対して単一の動きベクトルを指定するだけのブロックに基づく動き補償技術に関係したビジュアルアーチファクトの原因となる。

【発明の開示】

【発明が解決しようとする課題】

【0011】

本発明の目的はこれら従来技術の限界を克服することにある。

【0012】

より正確には、本発明の1つの目的は、画像データを符号化及び復号するための技術であって、ビットレートを小さくしかつ／または周知の技術で得られる再生画像のクォリティを改善することができるものを提供することにある。

【課題を解決するための手段】

【0013】

この目的は本発明による画像符号化方法によって達せられる。この画像符号化方法は、それぞれ様々な最適化基準に基づいてビデオシーケンスの少なくとも1つの画像のコンプレッション (compression) を最適化する少なくとも2つの方式の画像符号化を選択的に実行する。

【0014】

本発明のいくつかの有利な態様においては、復号器における前記符号化方式の中の1つの方式の選択に関する情報は、

- ・符号化と復号の段階において知られる、予め定められた選択と、
 - ・少なくとも一部の符号化された画像データを含むデータフローに含まれる前記選択の表示情報と、
 - ・符号化された画像データとは無関係なデータフローに含まれる前記選択の表示情報と、
 - ・前記復号器によって所定の方法で選択を判定すること、
- から成るグループに属する技術の中の少なくとも1つに基づいて復号器に知られてよい。

【0015】

有利には、前記画像符号化方法は、前記画像に適用すべき符号化方式を選択するための段階であって、

- ・画像の光度表現 (photometric representation) を最適化する第1の符号化と、
 - ・少なくとも2つの画像にわたる動きの表現を最適化する第2の符号化と、
- を少なくとも含む段階を具備する。

【0016】

こうして本発明は、特にMPEG符号化とメッシュ表現に基づく時間的補間を使用する符号化とのハイブリッド符号化による新規のビデオ符号化方法、並びに、その復号方法及びそれに関連するバイナリ表現の構造を提供する。

【0017】

好ましくは、前記第2の符号化は、前記第1の符号化を使用して符号化された少なくとも1つの以前の画像及び／または少なくとも1つの以後の画像を考慮する。

【0018】

有利には、前記第2の符号化は、前記第1の符号化を使用して符号化された直前の画像から計算される動きベクトル・フィールド及び／または前記第1の符号化を使用して符号化された直後の画像から計算される動きベクトル・フィールドを考慮に入れる。

【0019】

有利には、前記動きベクトル・フィールドはメッシュ網 (meshing) に適用される。

【0020】

10

20

30

40

50

この場合、前記動きベクトル・フィールドは、前記第2の符号化を使用して符号化された画像に関係した、推論された動きベクトル・フィールドを決定するために使用することができる。

【0021】

本発明の好ましい態様によれば、前記選択段階は固定因子Nのサブサンプリングの実施に基づいており、Nに1つの画像は前記第1の符号化を使用して符号化される。

【0022】

有利には、このNの値は少なくとも1つの所定の基準に応じて可変的である。

【0023】

或る特定の態様によれば、前記第1の符号化では画像ブロック上の変換とブロックによる時間的予測が実行される。この変換は例えばDCT型変換、アダマール（Hadamard）変換、ウェーブレット（wavelet）変換等である。 10

【0024】

画像ブロックは必ずしも正方形ではないが、必要性和利用可能な手段に応じてどんな形状も採ることができることがかわるであろう。

【0025】

この第1の符号化は、特に、MPEG-4またはH26L符号化であることが可能である。

【0026】

後者の場合、タイプI（intra）画像及び／またはタイプP（predictable）画像（好ましくはタイプB画像でない）を使用することが好ましい。 20

【0027】

本発明のもう一つの特別な特徴によれば、前記第2の符号化は有利には、Mを1以上であるとしてM階の例えば三角形型の階層型メッシュ網（hierarchical meshing）の使用に基づく。

【0028】

この場合、本方法は、好ましくは被覆領域（occlusion areas）またはブランキング領域（blanking areas）の管理段階（management phase）を具備する。

【0029】

作り出されたデータは、単一のフローにグルーピングされる。有利には、独立な伝送経路で伝送することができる少なくとも2つのデータフローを望むことができる。 30

【0030】

有利には、前記データフローは、グローバルヘッダと、

前記第1の符号化に基づいて符号化された画像データと、

前記第2の符号化に基づいて符号化された画像データと、を含むグループに属する。

【0031】

この結果、フロー伝送は独立に生じることが可能である。特に、このことによって、手段と必要性に応じて、画像を順繰りに及び／または部分的に復号することができるようになる。 40

【0032】

このため、本発明の特別な態様においては以下のような側面、すなわち、

- ・MPEG型またはITU-T/H.263型の標準化された符号化の構成モジュールによって推し進められる最適化と、

- ・メッシュ網に基づく技術の強力な時間的予測及び関係する誤り符号化と、を利用することができる。

【0033】

実際、メッシュ網に基づくアプローチは、連続的な動きフィールドを使用するおかげにより通常ブロック（normal block）の影響を避ける助けになる。さらに、メッシュ網技術によれば、ブランキング「オブジェクト」の検出、並びにこれらの領域に適切に適用される 50

誤り符号化が可能となる。その上、これらの領域周辺でのMPEG誤り符号化を組み合わせることによって、補間の効率はMPEGによって与えられる双方向画像（タイプB画像）のコストよりもずっと低いコストに十分改善されることが可能である。

【0034】

従って、基本的な情報は、MPEG型符号化を使用して低い時間解像度でも高いクォリティで効果的に符号化されることができる。そして、シーケンスのフロー全体は、メッシュ網による補間方式における符号化を使用して復元されることができる。

【0035】

本発明は、もちろん、

- ・前記符号化方法を使用して符号化された画像信号を復号するための方法と、 10
- ・前記符号化方法を使用して符号化された画像信号を符号化するための方法と、
- ・前記符号化方法を使用して符号化された画像信号を復号するための装置（有利には、符号化において実施されたものと類似した、ベクトル・フィールドの少なくとも一部及び／またはブランキング領域（blanking areas）の少なくとも一部を決定するための手段を含む）と、
- ・前記符号化方法を使用して符号化された少なくとも1つの画像信号を記憶するための装置と、
- ・前記符号化方法を使用して符号化された画像信号の符号化、伝送及び／または復号のシステム（有利には、前記符号化方式の1つの選択に関する情報は、 20
- ・符号化と復号の段階において知られる、予め定められた選択と、
- ・少なくとも一部の符号化された画像データを含むデータフローに含まれる前記選択の表示情報と、
- ・符号化された画像データとは無関係なデータフローに含まれる前記選択の表示情報と、
- ・復号器によって固有の方法で選択を判定すること、
- ・前記符号化方法を使用して符号化された画像信号を符号化及び／または復号するためのコンピュータ・プログラム製品と、
- ・そうしたプログラムのためのデータ支援と、から成るグループに属する技術の中の少なくとも1つに基づいて復号器に知られる）、も提供する。

【0036】

本発明は前記符号化方法に基づいて符号化されたデータを含む画像データ信号も提供する 30

【0037】

有利には、この信号は、前記符号化方法が作用したか否かを特定する少なくとも1つの指標を含む。

【0038】

好ましくは、前記信号は、ビデオシーケンスの先頭及び／またはそれぞれの信号フレームにフレーム構造を特定するデータを含む。

【0039】

有利には、前記第2の符号化を使用して符号化されたシーケンスは、この第2の符号化に基づいて符号化されたフレームの数を示すヘッダから始まる。 40

【0040】

或る特定の態様によれば、前記信号は、別個の伝送経路で伝送することができる少なくとも2つのデータフローを含む。

【0041】

この場合、前記データフローは有利には、

グローバルヘッダと、

前記第1の符号化に基づいて符号化された画像データと、

前記第2の符号化に基づいて符号化された画像データと、を含むグループに属する。

【0042】

本発明は、様々な領域、特に、 50

- ・デジタルTVと、
- ・IPネットワーク上のリアルタイム・ビデオと、
- ・移動体通信ネットワーク上のリアルタイム・ビデオと、
- ・画像データストレージと、を含むグループに属する領域に用途が見出される。

【0043】

本発明の他の特徴と利点は、本発明の以下の好ましい実施の態様の説明においてより明らかとなる。またその説明においては図面が参照される。ただし、あくまでもその説明は例示目的で行われるものであって、本発明はそこに与えられた例には限定されない。

【発明を実施するための最良の形態】

【0044】

本発明の好ましい実施態様を以下詳細に説明する。この態様は、例えばMPEG-4といったMPEG型符号化と、MPEG標準におけるB方式またはB画像とも呼ばれる補間方式で働くメッシュ符号化とのハイブリッドに基づく。

【0045】

ここに言及されたMPEG-4符号化は、均等な技術に基づく任意の符号器、すなわち、ブロック構造に基づく時間的予測及び離散コサイン変換と、生成された情報の定量化とエントロピー符号化を使用する任意の符号器で置き換えることができることに注意すべきである。特に、MPEG-4符号化をITU-T/H.263++符号化に置き換えることが可能である。

【0046】

符号器に入ってくるシーケンスの各画像毎に、その符号器はMPEG-4符号化モジュールまたはメッシュ網に基づく符号化モジュールを備えた符号器の或る特定の判定プロセス（例えば、固定因子型時間的サブサンプリング（fixed factor temporal subsampling））に基づいて判定を行う。

【0047】

メッシュ方式で符号化された画像は、それらの時間的予測の基準として、それらが属するメッシュ方式で符号化された画像のグループの直前または直後に置かれたMPEG-4方式で符号化された画像を使用する。

【0048】

本発明のコンプレッション効率のキーポイントは、メッシュ網に基づく動き補償によって、伴う符号化コストが低い非常に強力な時間的予測がもたらされる、ということである。

【0049】

実際、この技術は、

- ・画像内での様々なタイプの動きを考慮し、
- ・特に、オブジェクトの動きによる領域のカバリングとアンカバリングを処理する。

【0050】

本符号器の動作原理の概要を図1に示す。

【0051】

まず、入力画像は、或る与えられた判定方法に基づいて、例えばN個の内1個の画像がMPEG符号化され、その残りの画像はメッシュ網補間方式（meshing interpolated mode）で符号化されるといったように、予め定義されたリズムに従ってMPEG符号化モジュールまたはメッシュ網に基づく符号化モジュールのいずれかに送られる。 N_k はMPEG方式で符号化された画像の番号を表すことに注意する。

【0052】

その他全ての画像 I_l 、 $N_k < l < N_{k+1}$ は、B方式と呼ばれる補間方式（interpolated mode）で働く例えば三角形型のメッシュ網に基づく符号器によって符号化される。この符号器の一般動作原理は以下のようなものである。

【0053】

1. 画像 N_k と画像 N_{k+1} の間の前後の動きフィールドの計算。これらのフィールドは三角メッシュ網の形でモデリングされる。

10

20

30

40

50

【0054】

2 a. 予測可能領域の推定：補間の間に、 I_t 及び $I_{t'}$ の間の動きが推定される。動きベクトルの適用が対称変換 (conversion) を引き起こさない場合 (予測可能領域が存在することを意味する)、動き補償 (movement compensation) は $t+k$ が $[t, t']$ に属するとき、 I_{t+k} を補間するためにベクトルをスカラー k ($0 \leq k \leq 1$) で重み付けすることによって実行される。動き補償には、以下説明される3つの可能なアプローチが存在する。

【0055】

2 b. 非予測可能領域の推定：符号化される画像 I_1 における、これらの動きフィールドの知識からの、非予測可能なブランキング領域 (blanking areas) の検出。

10

【0056】

3. 以下の3つの可能な方式の1つに基づく、これらのブランキング領域の特定の符号化：

- ・動き補償が伴わない参照画像 (N_k 、 N_{k+1} 、またはそれらの動きフィールドを使って動き補償されたこれらの画像) の1つによる予測、次いで予測誤差を三角メッシュ網・ベース技術を使って符号化する。

- ・イントラ画像動き補償 (intra-image movement compensation) が伴う参照画像 (N_k 、 N_{k+1} 、またはそれらの動きフィールドを使って動き補償されたこれらの画像) の1つによる予測、次いで予測誤差を三角メッシュ網・ベース技術を使って符号化する。

20

- ・三角メッシュ網・ベース技術に基づくイントラ画像符号化 (intra-image encoding)。

【0057】

4. 選択的に、ブランキング領域周辺の領域に限定された、残りの予測誤差の P 方式 MPEG 型符号化または符号化。

【0058】

既に言及したように、動き補償は、事前推定 (estimate before)、事後推定 (estimate after)、あるいは事前推定及び事後推定の両方に基づく、3つのアプローチを使用して実行される。

【0059】

1) 事前推定による場合

補間の際、 I_{t1} と I_{t2} の間の動きが推定される。動きベクトルの適用が対称変換 (conversion) を引き起こさない場合 (予測可能領域に対応する)、動き補償は、 m が $[t1, t2]$ に属するとき I_m を補間するためにベクトルをスカラー $k = m / (t1 + t2)$ ($0 < k < 1$) で重み付けすることによって実行される。E1 推定画像が得られる。

30

【0060】

2) 事後推定による場合

補間の際、 I_{t2} と I_{t1} の間の動きが推定される。動きベクトルの適用が対称変換を引き起こさない場合 (予測可能領域に対応する)、動き補償は m が $[t1, t2]$ に属するとき I_m を補間するためにベクトルをスカラー $k' = 1 - m / (t1 + t2)$ ($0 \leq k' \leq 1$) で重み付けすることによって実行される。E2 推定画像が得られる。

40

【0061】

3) 事前推定と事後推定の両方による場合

補間の際、 I_{t1} と I_{t2} の間と I_{t2} と I_{t1} の間の動きが推定される。動きベクトルの適用が対称変換を引き起こさない場合 (予測可能領域に対応する)、動き補償は「事前」ベクトルをスカラー k ($0 \leq k \leq 1$) で重み付けすることによって実行され、E1 が得られる。同じことがスカラー k' ($0 \leq k' \leq 1$) を使って「事後」ベクトルについても行われ、E2 が得られる。このとき推定画像は、 a を ($0 \leq a \leq 1$) として、 $E = a E1 + (1 - a) E2$ である。

【0062】

2 ビット毎に、推定されたソリューションとそれに関するソース・イメージとの間で PSNR を計算した後にベストなソリューションが採られる。

50

【0063】

1. 画像 N_k 及び画像 N_{k+1} の間の前後の動きベクトル・フィールドの計算

画像 N_k 及び画像 N_{k+1} の間の前後の動きベクトル・フィールドが階層型メッシュの形で、例えば図2に示された三角分割 T_k^b と T_{k+1}^b 、の形で計算される。

【0064】

これらのメッシュ網は、或る決まったメッシュを分割することによって得られる。例えば、三角メッシュは、動き推定プロセスの或る特定の基準に基づいて4つの部分三角形に分割される。各階層レベルで、各メッシュに対して分割判定がなされたりなされなかったりする。これらの分割が決まると、均一なメッシュ構造を維持するために分割されたメッシュに隣接するメッシュが分割される。分割以前の初期のメッシュ（階層構造の頂点）はどんな形を採ることもできる。 10

【0065】

図2の例では、動き推定器（movement estimator）は三角形3と三角形8を分割する。この分割に伴って三角形2、4、7及び9が分割される。このプロセスは所定の階層レベルまで繰り返される。

【0066】

三角メッシュ網（triangular meshing）の場合、三角メッシュ網 T によって定義される動きフィールド表現（movement field representation）は各三角形 e に対し次式で与えられる。

【数1】

20

$$\vec{d}(\vec{p}, D) = \sum_{\{\text{ver}(e)\}_n} \psi^e(\vec{p}) \vec{d}_n \quad \forall \vec{p}(x, y) \in e$$

ここで、

- ・ e は x 、 y 座標の現ポイント $\vec{p}$ を含むメッシュ網 T の三角形要素を表す。
- ・ $\{\text{ver}(e)\}$ は図で i 、 j 、 k の番号が付けられたポジション $\vec{p}_i$ 、 $\vec{p}_j$ 、 $\vec{p}_k$ にある全ての3つのノード（結節点（node））またはピーク（頂点（peak））の集合を表す。
- ・ Ψ_l ($l = i, j, k$) は三角形要素 $e_{i,j,k}$ におけるポイント $\vec{p}(x, y)$ の重心座標を表し、次式で定義される。

30

【数2】

$$\left\{ \begin{array}{l} \Psi_l^e(x, y) = \alpha_l^e + \beta_l^e x + \gamma_l^e y, \alpha_l^e, \beta_l^e, \gamma_l^e \in \mathbb{R} \quad \text{if } \vec{p}(x, y) \in e_{i,j,k} \\ \sum_{l=i,j,k} \Psi_l^e(x, y) = 1 \\ \Psi_l^e(x, y) = 0 \quad \text{otherwise} \end{array} \right.$$

40

【0067】

このタイプのモデルは十分に連続的なフィールドを定義する。さらに、それは表現精度の十分なコントロール、コンプレッションには不可欠な特徴、を可能にする。

【0068】

メッシュ網階層構造（meshing hierarchy）の各レベルにおいて、予測誤差を最小化するためにノード動きベクトル（nodal movement vector）が計算される。メッシュ網に基づく様々な動き推定器を使用することが可能である。それらについては、例えば、仏国特許第98 11227号または仏国特許第99 15568号を参照されたい。

50

【0069】

思い出さなければならない重要なポイントは、最終メッシュは初期のメッシュからの分割による階層プロセスの結果である、ということである。この階層性は実際に或るノードとその親ノード（それが挿入されるアークの端）との間のノード動きベクトル (nodal movement vectors) の差分符号化 (differential encoding) に役立てるために使用される。メッシュ網構造は復号器においてメッシュ分割指標 (mesh division indicators) だけでなく初期のメッシュからの知識を使用して再計算される。

【0070】

この結果、プロセスの最後には、画像 N_k 及び画像 N_{k+1} から構成された画像のグループであってそのグループの全ての画像を再生するために画像 N_k 及び画像 N_{k+1} が使用されるそうしたグループのそれぞれに対して、2つの動きメッシュ網 (movement meshing) が得られる。

【0071】

2. ブランキング領域の検出

これら2つのメッシュ網から、オブジェクトのカバリングまたはアンカバリングによる、画像 N_{k+1} から得られる画像 N_k または画像 N_k から得られる画像 N_{k+1} における非予測可能領域であるブランキング領域 (blanking areas) が検出される。

【0072】

これらの領域は、ノード動きベクトルによって三角形が移動させられた後のそれらの三角形のカバリングによって単純に画定される。

【0073】

図は、移動後の三角形のアンカバリングの際の基本的なブランキング検出を示している。

【0074】

符号器は、偏りがより小さい動きベクトルを得るために、ブランキング領域における三角形を取り除くこと (deactivating) によって動き推定を継続させることができる。

【0075】

それでもなお、これは厳密には符号器の戦略内のことであって、最後に、符号化されバイナリ・ストリームに挿入されるのは2つの完全な動きメッシュ網 T_k^b と T_{k+1}^f である。復号器はこれら2つのメッシュ網を使ってブランキング領域を見つけることができる。

【0076】

これらのブランキング領域は画像 N_k 及び画像 N_{k+1} において画定され、それらが検出されるとすぐに、それらに対応する三角形は結果として符号器と復号器においてラベル付けされる。

【0077】

それでもなお、符号器は画像 N_{k+1} 及び画像 $N_{k+1} - 1$ のどこにこれらの領域が見つかるかを知る必要がある。これらは、現画像と参照画像 N_k または N_{k+1} との間の時間的距離を考慮するよう再規格化されたノード動きベクトルを適用することによって符号化される画像の上にメッシュ網 T_k^b と T_{k+1}^f を投影することによって単純に得られる。

【0078】

3. ブランキング領域の符号化

各ブランキング領域に対して、可能な予測の参照画像は I_{Nk} 、 $I_{N(k+1)}$ 、並びに I_{c_1} から選ばれる。 I_{c_1} はメッシュ・カバリングがまだ存在しないレベルにおいて、メッシュ網 T_k^b と T_{k+1}^f での動き補償によって、1において得られる画像である。

【0079】

より正確には、 I_{Nk} と $I_{N(k+1)}$ の間の選択は単純に、現ブランキング領域を生成したメッシュ網 T_k^b または T_{k+1}^f に依存する。次いで、このメッシュ網は画像を予測して I_{c_1} を与えるために使用される。 I_{c_1} と I_{Nk} または $I_{N(k+1)}$ の間の選択は、最小誤りを引き起こす画像が維持されるという、予測誤差基準に基づく。従って、維持された予測の選択を符号化することは、領域毎に、バイナリ・フローに1ビットを挿入するという問題である。

【0080】

10

20

30

40

50

選択した参照画像に I は採らない。

これらの領域の符号化の残りは、以下の 2 つの段階、すなわち、

- ・予測段階と、
 - ・予測誤差、または不正な予測の場合における原テクスチャの符号化段階と、
- を含む。

【0081】

3. 1 ブランキング領域テクスチャの残りの予測

3 つの方法が排他式に使用することができる。判定は最低誤りの基準に基づく。

【0082】

方法 1

領域内のピクセルの値 Y 、 U と V は単純に、同じ参照画像 I_1 内に置かれたピクセルの値である。 $\hat{I}_l$ を結果の画像とする。次に $\hat{I}$ と I_1 の間の予測誤差が符号化される。

10

【0083】

方法 2

動き推定がブランキング領域上で I_1 (符号化される画像) と $\hat{I}_l$ (方法 1 からの予測結果) の間で実行される。最後のメッシュ網レベル T_r 、 $r = k$ または $k + 1$ 、からメッシュがカバーされる前に生じる結果のメッシュは、ノードの動きと同様に符号化される。最後に、残りの予測誤差は後に定義される手続きに基づいて符号化される。

20

【0084】

方法 3

予測は一切実行されず、領域内のピクセルの最初の値が符号化される。

【0085】

4. ブランキング領域のテクスチャまたは予測誤差の符号化

原テクスチャと予測誤差は同じ符号化を経験する。その原理を以下に述べる。

【0086】

初期の三角メッシュ網は符号化されるべき領域の予測のために維持される T_r メッシュ網の動きから定義することが可能であったと仮定する。この初期のメッシュ網を得る方法については以降述べられる。

30

【0087】

従って、テクスチャは、各メッシュ毎に以下の選択に基づいて近似される。

- ・多くの高い周波数を持つメッシュは DCT と呼ばれる離散コサイン変換に基づいて符号化される。
- ・より滑らかなメッシュは、仕上がった要素 (finished elements) の微細モデルを使用して符号化される。

【0088】

再び、メッシュ網を通じて表現を符号化するコストを減らすために階層的アプローチが使用される。

40

【0089】

使用されるアプローチは、維持されるべきメッシュの規則的階層に関係する符号化の低コストを可能にするが、その一方でメッシュの不規則な分割によって許される画像コンテンツの局所的な適用が許される。

【0090】

領域の密な初期のメッシュ網から、メッシュは或る与えられたレベルまで 4 つの三角部分メッシュに再分割される。

【0091】

最後のレベルでは、2 つの隣接する三角形から成る四角形の対角線を置換することによ

50

て近似誤差が減少する場合には、そうした置換を選択的に実行することができる。

【0092】

4. 1 ブランキング領域上でのテクスチャ・メッシュ網の初期化

このメッシュ網は単純に、議論している領域に対称変換 (conversion) が現れる前に最後のレベル T_r (維持される方向に応じた T_k または T_{k+1} の動きの結果もたらされるメッシュ網) によって与えられる。結果、動きメッシュ網に自然に挿入されるテクスチャ・メッシュ網が得られる。というのは、テクスチャ・メッシュ網は動きメッシュ網から取り出されるからである。

【0093】

4. 2 三角形上のテクスチャに使用される表現

2つの表現、すなわちアファイン補間 (affine interpolation) 及び三角形型 4 D C T (triangular 4 DCT)、が組み合わされる。

【0094】

アファイン補間 (affine interpolation)

三角メッシュ網のノードは光度 (色、エラー) 情報 (photometric information) を含み、三角形内部のポイントに対する補間はアファイン補間とも呼ばれるラグランジェ仕上要素 (Lagrange finished element) によって実行される。

【0095】

3つのノード $\bar{p}_l$ 、 $l=i, j, k$ 、によって定義される三角形 $e_{i,j,k}$ 内部のポイント $\bar{p}(x, y)$ の値 $v(\bar{p})$ は次式で与えられる。

【数 3】

$$v(\bar{p}) = \sum_{l=i,j,k} \psi_l^e(x, y) \bar{p}_l$$

ここで、 Ψ_l ($l=i, j, k$) はポイントの重心座標を表す。

【0096】

$v(\bar{p})$ はそのポイントの光度成分 Y、U 及び V の 1 つまたはこれら 3 つの成分の予測誤差であることが可能である。

【0097】

ノードの値を形算するためにいくつかの方法、特に最小自乗法 (method of lesser squares)、を使用することができる。

【0098】

三角形上の離散コサイン変換 (DTC)

この方法の原理は、任意の三角形を直角二等辺三角形に変換することにある。つまり、この三角形の中身は、対称正方行列を与えるために斜辺に関して対称的にされる (図 4)。

【0099】

標準的な DTC (正方形) は従ってこの行列に適用される。変換された行列も対称的であることを示すことができる。従って下側の三角形の係数だけが定量化され、統計的に符号化 (エントロピー符号化) される。

【0100】

図 4 は、様々な段階のプロセスを記述する。具体的には、三角形 T を選び、その三角形 T を直角二等辺三角形 T' にアファイン変換する。アファイン変換が行われると、三角形のピクセルはもはや正規直交格子になく、基準三角形内部の光度値の再サンプリングが推奨される。そうするために、画像における動き補償のためのプロセスに似たプロセスが (アファイン変換の際に)、インターポレータ (interpolator)、例えば双線形 (bilinear

10

20

30

40

50

）インターポレータ、を使用して使用される。

【0101】

アファイン変換 F とその逆変換 F^{-1} は次式によって定義される。

【数4】

$$F: \begin{cases} X = \frac{(x_2 - x_1)(y_1 - y) + (y_1 - y_2)(x_1 - x)}{(x_3 - x_1) + (y_2 - y_1) + (y_1 - y_3)(x_2 - x_1)} N \\ Y = \frac{(x_1 - x_3) + (y_1 - y) + (y_3 - y_1)(x_1 - x)}{(x_3 - x_1)(y_2 - y_1) + (y_1 - y_3)(x_2 - x_1)} N \end{cases}$$

10

$$F^{-1}: \begin{cases} x = x_1 + (x_3 - x_1) \frac{X}{N} + (x_2 - x_1) \frac{Y}{N} \\ y = y_1 + (y_3 - y_1) \frac{X}{N} + (y_2 - y_1) \frac{Y}{N} \end{cases}$$

【0102】

三角形 T' （つまり T' の対称ブロック M ）の光度値 $M(i, j)$ は、逆変換 F^{-1} と、次いで符号化する画像のテクスチャの Γ 補間とによって次のように得られる。

【数5】

20

$$M(i_Y, j_X) = I_{\Gamma}(P_{(x,y)}) = I_{\Gamma}(F^{-1}(Q_{(x,y)}))$$

ここで、係数 $M(i_Y, j_X)$ は、 $P(X, Y) = F^{-1}(Q)$ なるポイント $Q(X, Y)$ における値である。

【0103】

I_{Γ} は、そのポイントにおける画像の値を計算するために使用されるインターポレータを表す。そのポイントの座標は非整数の場合がある。

【0104】

30

Γ テクスチャの再現 $\hat{\Gamma}$ は次式で与えられる。

【数6】

$$\hat{\Gamma}(P_{(x,y)}) = I_{\Gamma}(Q_{(x,y)}) = I_{\Gamma}(F(P_{(x,y)}))$$

ここで、 I_{Γ} は M の定量化されたバージョンである M' ブロック値から補間されたテクスチャを表す。

40

【0105】

この技術は、非ゼロ領域 (non-null area) の三角形に適用できるだけである。しかし、定義により、このタイプの三角形はテクスチャ符号化を必要としない。

【0106】

SADCT（形式 (form) に適用される DCT）とは対照的に、この変換は、たとえ定量化がなくても、逆変換後の完全な再生を保証しない。

【0107】

再生誤差を減らすために、三角形 i の M_i ブロック（サイズは $N_i \times N_i$ ）を計算するためにスケール因子 α が導入される。

【数7】

50

$$N_i = E(\alpha\sqrt{2A_i})$$

ここで、

- ・ Eは超過による完全な部分、
- ・ A_i は三角形 i の面積。

【0108】

実際、 $\alpha = 1$ は、二等辺に近い三角形にはより効果的な興味ある妥協案を提供する。 $\alpha < 1$ のケースは、情報量をコンプレッションするために定量化ステップと併せて使用される 10

【0109】

ブロック M_i が各三角形に定義されると、標準的な D C T 変換がそれらに適用され、変換された係数はいくつかの可能な手続き、例えば一様なスカラー定量化または変換された係数の周波数に伴って増加する定量化に基づいて定量化される。よく知られた M P E G 符号化または J P E G 定量化行列も可能である。

【0110】

D C T 表現は次式によって与えられる。

【数 8】

$$F(u, v) = \frac{1}{\sqrt{NM}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \lambda(u) \lambda(v) \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2M} f(i, j)$$

$$\lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \xi=0 \\ 1 & \text{otherwise} \end{cases}$$

20

【0111】

定義により $f(i, j) = f(j, i)$ ($\forall_{u,v,i,j} = 0, \dots, N-1$) であるので、全ての u, j に 30
対して $F(u, v) = F(v, u)$ の関係が存在する。

結果、変換された行列の下側の部分の係数を計算するだけで十分である。

【0112】

4. 3 テクスチャのグローバルな符号化

既に述べたように、或る与えられた階層レベルに属する三角形をそれぞれ 4 つの部分三角形に分割して、アーク (arc) の中点にノードを挿入することによって得られる均一な階層型メッシュ (hierarchical meshing) が使用される。このプロセスは最大レベルに達するまで繰り返される。三角形のこの階層構造は、4 木 (quaternary tree) 形式で表現され、符号器によって管理される (図 5 参照)。符号化すべき領域に含まれる三角形だけが 40
考慮されることに注意する。基本的な初期のメッシュ網の構築手続きは、メッシュ階層構造の全ての三角形が符号化すべき領域に属することを保証する。

【0113】

ブランキング領域をメッシングすることによる符号化プロセスは以下のようにまとめられる。

【0114】

1. 鱗状 (interleaved) の階層型メッシュ網は、正規の初期のメッシュ網を生成して、次に各三角形を 4 つの部分三角形に再分割して、アークの中点に新たなノードを挿入することによって、符号化すべき領域上に定義される。ノード値は、メッシュ網による面積の近似誤差を最小化するよう計算される。

50

【0115】

2. ピクセル値は、そのノード値を使用してそのピクセルを含む三角形上でアファイン補間を使用していっしょに近づくように動かされる。

【0116】

階層構造の三角形それぞれに対して、近似誤差 E が評価され、次いで、様々な表現及び符号化の方法が2つの閾値 σ_1 、 σ_2 に基づいて判定される。

1. $E < \sigma_1$ の場合、三角形上のアファイン補間は十分である。

2. $\sigma_1 < E < \sigma_2$ の場合、三角形の最も細かい分割がアファイン補間を常に使用して良い近似を得るために使用されなければならない。

3. $E > \sigma_2$ の場合、三角形がテクスチャ（絵柄）化され、アファイン補間誤差が DCT を使用して符号化される。 10

【0117】

最後に、最も細かいメッシュ網において、2つの隣接する三角形によって形成された四角形の対角線を置換することによって生み出される誤差の減少が調べられる。肯定的な結果が得られれば、この置換は有効となる。

【0118】

異なる三角形に対して選ばれた符号化方法に基づいて、様々なデータが以下の仕方で符号化される。

【0119】

最初に、YUV ノード値が親ノード（最新のノードが挿入されるアークの端）の値を使用して予測される。こうしてノード値とその予測値との間の差が定量化される。 20

【0120】

最後に、4木構造（三角形の分割と非分割の指標（indicator）を含む）、対角線置換指標、YUV の差分ノード値と定量化された DCT 係数は算術符号化（arithmetic encoding）を使用して符号化され、バイナリ・ストリームに挿入される。

【0121】

5. メッシュ符号化されたフレームのバイナリ・フローにおける符号化された情報のまとめ

N_{k+1} と $N_{k+1} - 1$ （ N_k と N_{k+1} はそれぞれ、MPEG 方式で符号化された前と後のフレームを表す）の間のメッシュ方式で符号化されたフレームの各グループはバイナリ・フローの中身全部として表現される。 30

【0122】

運ばれるデータ（driven data）は、符号化された形で、以下のものを含む。

フレーム・グループのセットに対する、特に符号化されたフレームの実数を含む、ヘッダ。

動きメッシュ網（ノードの構造及び動きベクトル） T_k^b と T_{k+1}^f 。

グループの各画像に対する、予測またはオリジナルのエラー・テクスチャ。

【0123】

6. バイナリ・ストリームのグローバル構造

グローバルなバイナリ・ストリームは、図8にあるように、MPEG 符号化された一連のフレームとメッシュ符号化されたフレームのグループから構成される。 40

【0124】

符号化されたシーケンスを表すバイナリ・ストリームのグローバル・ヘッダは、特に、ハイブリッド符号化（hybrid encoding）の指示を含む。

【0125】

メッシュ符号化されたフレームのグループの対応するバイナリ・ストリームの部分は、中でも実際に符号化されたフレームの数を示し、ゼロの場合があるヘッダから始まる。

【0126】

符号化されたシーケンスのグローバル・ヘッダや、MPEG 符号化された画像と補間メッシュ方式で符号化された i 画像のグループにそれぞれ対応する様々なデータフロー（バイ 50

ナリ・ストリーム)は、必要な場合には、異なる独立な伝送路で送られることが可能である。特に、この符号化手続きは、シーケンスの階層型(またはスケーラブルな)符号化、すなわち、全ビットレートの一部を使用するだけの復号、を許す。

【0127】

7. 復号手続き

図9に復号原理の概要を示す。

【0128】

最初に、ヘッダの復号によってハイブリッド復号の開始が可能となる。

【0129】

次に、復号器は、独立な構成要素に対応するバイナリ・ストリームの各部分に対して、それがMPEG符号化されたフレームかあるいはメッシュ符号化されたフレームのグループかを認識する。 10

【0130】

MPEG-4フレームは、MPEG-4復号モジュールに送られ、そしてメッシュ符号化されたフレームのグループは、メッシュ網に基づく復号モジュールに提供される。

【0131】

7.1 メッシュベース復号

最初に、画像グループ I_1 、 $N_k < l < N_{k+1}$ 、に対する動きメッシュ網 T_k^b と $T_{(k+1)}^f$ が復号される。

【0132】

次に、これらの画像のブランキング領域が符号器における手続きと同じ手続きに基づいて見出される。 20

【0133】

その後、そのブランキング領域の外側のピクセルが画像 I_{N_k} 及び $I_{N_{k+1}}$ と動きフィールド T_k^b と $T_{(k+1)}^f$ を使用して単純に補間される。

【0134】

最も粗いテクスチャのメッシュ網(階層構造の頂上)が符号器における手続きと同じ手続きに基づいてそれぞれのブランキング領域に対して見出される。

【0135】

こうして、対応する階層型メッシュ網に関する情報(三角分割指標、アファイン補間判定またはDCT符号化、差分ノード値YUVと定量化されたDCT係数)が復号され、これらの領域のピクセル値YUVが再生される。 30

【図面の簡単な説明】

【0136】

【図1】本発明の符号化方法の手続きの流れを示した概略図である。

【図2】動き(movement)に関する階層型メッシュ網構造の一例を示した図である。

【図3】三角メッシュにおけるアファイン補間原理を説明するための図である。

【図4】三角形のカバリングによって検出されるブランキング(blanking)の一例を示した図である。

【図5】対称正方行列の中の任意の三角形の変換過程を示した図である。 40

【図6】任意の三角形の直角二等辺三角形への変換過程を示した図である。

【図7】階層型メッシュ網とそれに関する4木表現を示した図である。

【図8】階層型メッシュ網の符号化判定の一例を示した図である。

【図9】本発明に基づくバイナリ・ストリームのグローバル構造を示した図である。

【図10】本発明に基づく復号器のブロック図である。

【図 1】

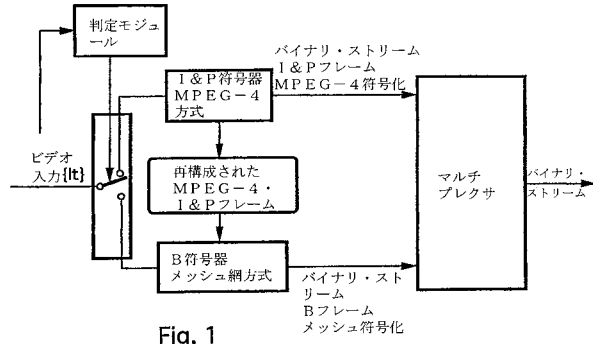


Fig. 1

【図 2】

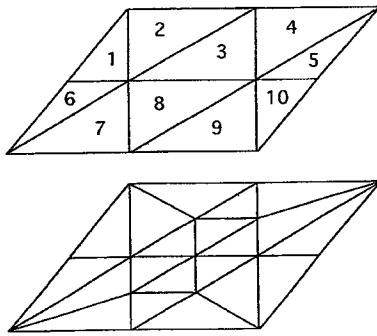


Fig. 2

【図 3】

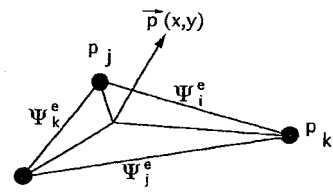


Fig. 3

【図 4】

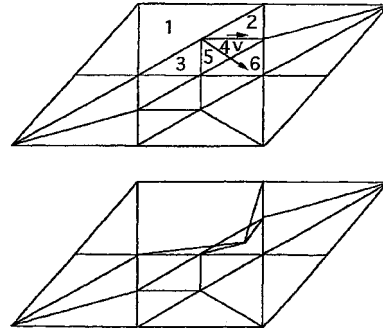


Fig. 4

【図 5】

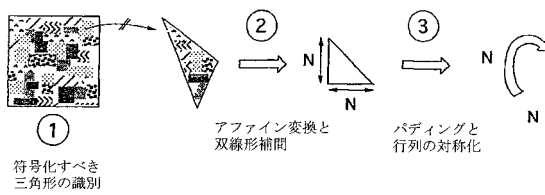


Fig. 5

【図 7】

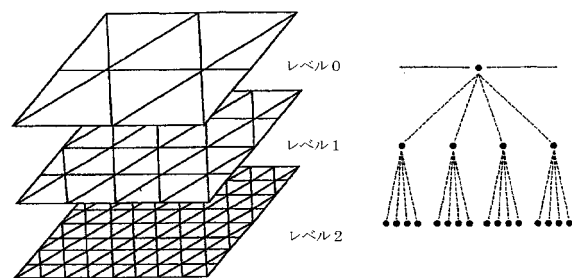


Fig. 7

【図 6】

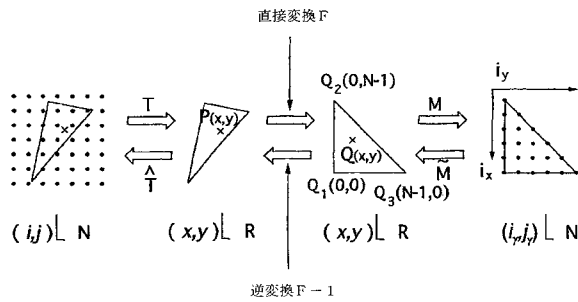


Fig. 6

【図 8】

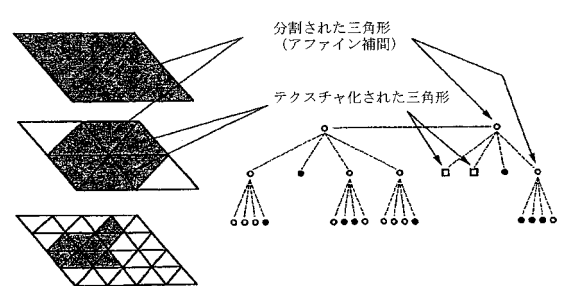


Fig. 8

【図 9】

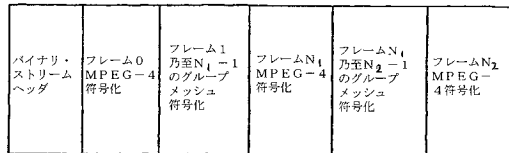


Fig. 9

【図 10】

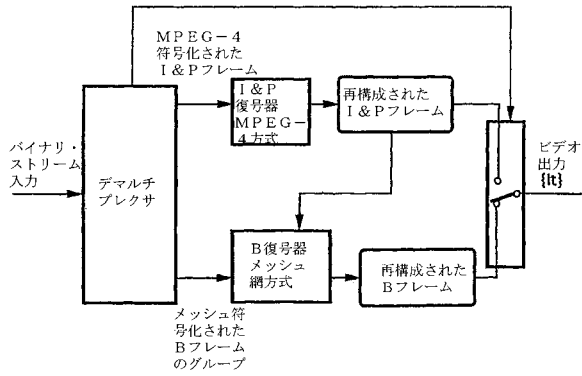


Fig. 10

【国際公開パンフレット】

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international(43) Date de la publication internationale
1 août 2002 (01.08.2002)

PCT

(10) Numéro de publication internationale
WO 02/060184 A1(51) Classification internationale des brevets⁷ : H04N 7/46,
7/36(21) Numéro de la demande internationale :
PCT/FR02/00322(22) Date de dépôt international :
25 janvier 2002 (25.01.2002)

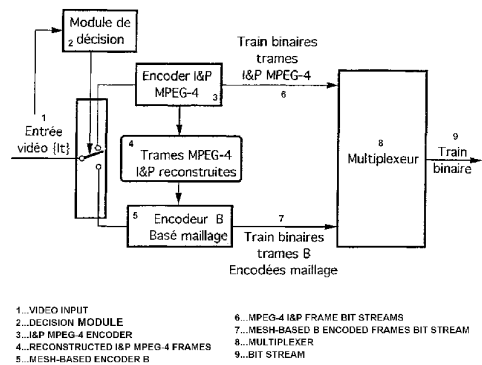
(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
01/01174 26 janvier 2001 (26.01.2001) FR(71) Déposant (pour tous les États désignés sauf US) :
FRANCE TELECOM [FR/FR]; 6 Place d'Allemey,
F-75015 PARIS (FR).(72) Inventeurs; et
(75) Inventeurs/Déposants (pour US seulement) : SANSON,
Henri [FR/FR]; 12 Allée des Roscaux, F-35090 ACIGNIE
(FR); LAURENT-CHATENET, Nathalie [FR/FR]; 3 rue
des Fraiches, F-35630 VIGNOC (FR); BUISSON, Alexan-
dre [FR/FR]; Résidence "Ondine", Appart B211, 18 rue
Noël Blayay, F-35000 RENNES (FR).(74) Mandataire : VIDON, Patrice; Cabinet Patrice Vidon,
Le Nobel - Bât. A, 2 Allée Antoine Becquerel, BP 90333,
F-35703 Rennes Cedex 7 (FR).(81) États désignés (national) : AU, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SG, SI,

[Suite sur la page suivante]

(54) Title: IMAGE CODING AND DECODING METHOD, CORRESPONDING DEVICES AND APPLICATIONS

(54) Titre : PROCÉDE DE CODAGE ET DE DECODAGE D'IMAGES, DISPOSITIFS ET APPLICATIONS CORRESPON-
DANTS(57) Abstract: The invention concerns a method for coding images using selectively at least two image coding modes, each opti-
mising compression of at least a video sequence image on the basis of different optimisation criteria.

[Suite sur la page suivante]



WO 02/060184 A1

WO 02/060184 A1



SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

Publiée :

- avec rapport de recherche internationale
- avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues

(84) **États désignés (régional) :** brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CL, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du FCT.

(57) **Abstrégé :** L'invention concerne un procédé de codage d'images mettant en oeuvre sélectivement au moins deux modes de codage d'images, optimisant chacun la compression d'au moins une image d'une séquence vidéo en fonction de critères d'optimisation différents.

WO 02/060184

PCT/FR02/00322

PROCÉDÉ DE ET DE DÉCODAGE D'IMAGES, DISPOSITIFS,
ET APPLICATIONS CORRESPONDANTS

Le domaine technique de l'invention est celui du codage à réduction de débit des séquences d'images animées, vidéo notamment.

5 Les applications du codage vidéo sont particulièrement nombreuses.
Citons de manière non exclusive :

- La transmission de la TV numérique ;
- La transmission vidéo en temps réel sur différents types de réseaux :
IP, mobiles, ("Streaming IP") ;
- 10 - Le stockage informatique de vidéos.

L'invention s'applique notamment dans le cadre de systèmes mettant en œuvre un codage de type MPEG. Par codage de type MPEG, on entend un codage basé sur une prédiction temporelle et une transformation en cosinus discrète basée sur une structure de blocs rigides, souvent de taille fixe, mais éventuellement de
15 taille variable. Les 2 standards représentatifs pour cette famille de codages sont les normes MPEG-4 de la version 1 à 4 et ITU-T/H.263 jusqu'à la version 2. L'invention peut également s'appliquer dans le cadre de la recommandation du CCITT H26L (voir par exemple le document VCEG-N83d1 correspondant).

20 Les schémas de codage et décodage vidéo proposés à ce jour se répartissent en 2 catégories :

- Les codages normalisés, soit par l'ISO/MPEG, soit par l'ITU-T, tous basés sur le même type de techniques (prédiction temporelle et une transformation en cosinus discrète basées sur une structure de blocs)
- Les codages en cours de développement proposés par des laboratoires
25 de recherche, qui font appel à un large panel de techniques : Codage par Ondelettes, par Régions, par Fractales, par Maillages, etc. ..

Actuellement, le codage MPEG-4 est considéré comme l'état de l'art non seulement des codages normalisés, mais aussi de l'ensemble des codages publiés.

30 Les codages de type MPEG-4 ou ITU-T/H.263++ sont considérés comme ayant atteint leurs limites, notamment à cause de la structure de blocs rigides de

taille fixe utilisée comme support de l'ensemble des calculs et opérations de codage. En particulier, la prédiction temporelle des images au sein d'une séquence est insuffisamment exploitée.

5 Par ailleurs, les codages alternatifs publiés n'ont pas encore atteint un degré suffisant d'optimisation.

Ainsi, pour permettre d'obtenir des séquences vidéos codées à bas débit, les codeurs procèdent en général à une réduction de la taille des images et au sous-échantillonnage temporel de la séquence vidéo originale. Toutefois, la seconde technique a pour inconvénient de restituer des mouvements saccadés plus ou
10 moins gênants pour l'utilisateur en fonction du niveau de sous-échantillonnage.

Pour éviter de telles saccades, il importe de reconstruire les images manquantes (non codées) au décodeur par interpolation temporelle.

Cependant, les techniques actuelles d'interpolation temporelle d'images ne permettent pas d'obtenir de résultats satisfaisants, surtout lorsqu'elles sont mises
15 en œuvre au seul décodeur. En effet, ces techniques sont la source d'artéfacts visuels liés aux techniques de compensation de mouvement basées blocs qui ne définissent qu'un seul vecteur mouvement pour l'ensemble des pixels d'un bloc.

L'objectif de l'invention est précisément de pallier les limitations des techniques antérieures.

20 Plus précisément, un objectif de l'invention est de fournir une technique de codage, et de décodage, de données d'images, qui permette d'obtenir un débit réduit et/ou une meilleure qualité d'image reconstruite, par rapport aux techniques connues.

Cet objectif est atteint, selon l'invention, à l'aide d'un procédé de codage
25 d'images, mettant en œuvre sélectivement au moins deux modes de codage d'images, optimisant chacun la compression d'au moins une image d'une séquence vidéo en fonction de critères d'optimisation différents.

Selon plusieurs modes de réalisation avantageux, une information sur le choix d'un desdits modes de codage peut être connue d'un décodeur selon au
30 moins une des techniques appartenant au groupe comprenant :

- choix prédéfini, connu au codage et au décodage ;
- information représentative du choix incluse dans un flux de données comprenant au moins certaines des données d'images codées ;
- 5 - information représentative du choix incluse dans un flux de données indépendant des données d'images codées ;
- détermination du choix de façon intrinsèque, par le décodeur.

Avantageusement, le procédé comprend une étape de sélection d'un mode de codage à appliquer à ladite image, parmi au moins :

- 10 - un premier codage optimisant une représentation photométrique d'une image ;
- un second codage optimisant une représentation du mouvement entre au moins deux images.

15 Ainsi, la présente invention concerne un nouveau procédé de codage vidéo par hybridation d'un codage notamment de type MPEG et d'un codage par interpolation temporelle basée sur une représentation par maillages, ainsi que le procédé de décodage et la structure de la représentation binaire associée.

Préférentiellement, ledit second codage tient compte d'au moins une image précédente et/ou au moins une image suivante codée à l'aide dudit premier codage.

20 De façon avantageuse, ledit second codage tient compte d'un champ de vecteurs mouvement calculé à partir de l'image immédiatement précédente codée à l'aide dudit premier codage et/ou d'un champ de vecteurs mouvement calculé à partir de l'image immédiatement suivante codée à l'aide dudit premier codage.

25 Avantageusement, ces champs de vecteurs sont appliqués à un maillage.

Dans ce cas, lesdits champs de vecteurs mouvement peuvent être utilisés pour déterminer un champ de vecteurs mouvement déduit, associé à une image codée à l'aide dudit second codage.

30 Selon un mode de réalisation préférentiel de l'invention, ladite étape de sélection repose sur la mise en œuvre d'un sous-échantillonnage de facteur fixé N,

une image sur N étant codée à l'aide dudit premier codage.

Avantageusement, cette valeur N est variable, en fonction d'au moins un critère prédéterminé.

5 Selon un mode de réalisation particulier, ledit premier codage met en œuvre une transformation sur des blocs d'images et une prédiction temporelle par blocs. Cette transformation est par exemple du type DCT, transformation de Hadamard, transformation en ondelettes,...

On notera que les blocs d'images ne sont pas forcément carrés, mais peuvent prendre toute forme adaptée aux besoins et aux moyens disponibles.

10 Ce premier codage peut notamment être d'un codage MPEG-4 ou H26L.

Dans ce dernier cas, on utilisera préférentiellement des images de type I (intra) et/ou de type P (prédictif) (et non, préférentiellement, des images de type B).

15 Selon un autre aspect particulier de l'invention, ledit second codage repose avantageusement sur la mise en œuvre d'un maillage hiérarchique de M niveaux, M étant supérieur ou égal à 1, et par exemple un maillage triangulaire.

Dans ce cas, le procédé comprend préférentiellement une étape de gestion des zones d'occlusions.

20 Les données produites peuvent être regroupées dans un flux unique. Avantageusement, on peut prévoir au moins deux flux de données, pouvant être transmis sur des voies de transmission indépendantes.

Lesdits flux de données appartiennent avantageusement au groupe comprenant :

- 25 - une entête globale ;
- des données d'image codées selon ledit premier codage ;
- des données d'image codées selon ledit second codage.

La transmission des flux peut donc avoir lieu de façon indépendante. Cela permet notamment un décodage progressif et/ou partiel des images, en fonction des moyens et des besoins.

30 Ainsi, selon un mode de réalisation particulier de l'invention, on exploite

notamment les aspects suivants :

- Optimisation poussée des modules constitutifs des codages normalisés de type MPEG ou ITU-T/H.263
- Prédiction temporelle et codage d'erreur associé performants pour les techniques basées maillage.

En effet, l'approche basée maillage permet d'éviter les effets de bloc habituels grâce à l'utilisation de champs de mouvement continus. De plus, la technique des maillages permet de détecter les occultations « d'objets », ainsi qu'un codage de l'erreur bien adapté à ces zones. En combinant de plus un codage d'erreur de type MPEG autour de ces zones, il est possible d'améliorer sensiblement l'efficacité de l'interpolation à un coût largement inférieur aux images bidirectionnelles (images de type B) proposées par MPEG.

Ainsi, on peut coder efficacement l'information de base à une faible résolution temporelle grâce au codage de type MPEG, avec une bonne qualité, puis restituer la pleine fluidité de la séquence grâce au codage en mode interpolé par maillage.

L'invention concerne également, bien sûr :

- les procédés de décodage d'un signal d'images codé à l'aide du procédé de codage décrit ci-dessus ;
- les dispositifs de codage d'un signal d'images codé à l'aide du procédé de codage décrit ci-dessus ;
- les dispositifs de décodage d'un signal d'images codé à l'aide du procédé de codage décrit ci-dessus (comportant avantageusement des moyens de détermination d'au moins une partie d'un champ de vecteurs et/ou d'au moins une partie des zones d'occlusion, similaires à ceux mis en œuvre lors du codage) ;
- les dispositifs de stockage d'au moins un signal d'images codé à l'aide du procédé de codage décrit ci-dessus ;
- les systèmes de codage, de transmission et/ou de décodage d'un signal d'images codé à l'aide du procédé de codage décrit ci-dessus

(le choix d'un desdits modes de codage pouvant avantageusement être connue d'un décodeur selon au moins une des techniques appartenant au groupe comprenant :

- choix prédéfini, connu au codage et au décodage ;
- 5 - information représentative du choix incluse dans un flux de données comprenant au moins certaines des données d'images codées ;
- information représentative du choix incluse dans un flux de données indépendant des données d'images codées ;
- 10 - détermination du choix de façon intrinsèque, par le décodeur ;
- les produits programme d'ordinateur pour le codage et/ou le décodage d'un signal d'images codé à l'aide du procédé de codage ;
- les supports de données d'un tel programme.

15 L'invention concerne également les signaux de données d'images comprenant des données codées selon le procédé décrit ci-dessus.

Avantageusement, ce signal comprend au moins un indicateur indiquant si le procédé est ou non activé.

20 De façon préférentielle, le signal comprend une donnée précisant la structure des trames, en début de séquence vidéo et/ou dans chaque trame de signal.

Avantageusement, une séquence codée à l'aide dudit second codage débute par une entête précisant le nombre de trames codées selon ce second codage.

25 Selon un mode de réalisation particulier, le signal comprend au moins deux flux de données, pouvant être transmis sur des voies de transmission indépendantes.

Dans ce cas, lesdits flux de données appartiennent avantageusement au groupe comprenant :

- une entête globale ;
- 30 - des données d'image codées selon ledit premier codage ;

WO 02/060184

7

PCT/FR02/00322

- des données d'image codées selon ledit second codage.

L'invention trouve des applications dans de nombreux domaines, et notamment dans les domaines appartenant au groupe comprenant :

- la télévision numérique ;
- 5 - la vidéo temps réel sur réseau IP ;
- la vidéo temps réel sur réseau vers les mobiles ;
- le stockage de données d'images.

D'autres caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description d'un mode de réalisation préférentiel de l'invention, donné à titre de simple exemple illustratif et non limitatif, et des
10 dessins annexés parmi lesquels :

- la figure 1 est un schéma de principe du codage de l'invention ;
- la figure 2 présente un exemple de structure de maillage hiérarchique pour le mouvement ;
- 15 - la figure 3 illustre le principe de l'interpolation affine sur une maille triangulaire ;
- la figure 4 est un exemple d'occultation détectée par recouvrement de triangles ;
- la figure 5 illustre le processus de transformation d'un triangle
20 quelconque de l'image en une matrice carrée symétrique ;
- la figure 6 illustre la transformation d'un triangle quelconque en un triangle isocèle rectangle ;
- la figure 7 illustre un maillage hiérarchique et la représentation par arbre quaternaire associée ;
- 25 - la figure 8 est un exemple de décision de codage pour le maillage hiérarchique ;
- la figure 9 présente une structure globale d'un train binaire selon l'invention ;
- la figure 10 présente un diagramme bloc d'un décodeur selon
30 l'invention.

Le mode de réalisation de l'invention décrit ci-après consiste essentiellement en l'hybridation d'un codage de type MPEG, par exemple MPEG-4 avec un codage par maillage fonctionnant en mode interpolé, appelé également mode B ou images B dans les normes MPEG.

5 Notons que le codage MPEG-4 mentionné ici peut être remplacé par tout codeur basé sur des techniques équivalentes, c'est-à-dire utilisant une prédiction temporelle et une transformation en cosinus discrète basées sur une structure de blocs, et les quantifications et codages entropiques pour les informations générées. En particulier, un codage ITU-T/H.263++ peut être substitué au codage MPEG-4.

10 Pour chaque image de la séquence entrant dans le codeur, ce dernier décide en fonction d'un certain processus de décision (par exemple, un sous-échantillonnage temporel de facteur fixé) de la coder avec le module d'encodage MPEG-4, ou bien avec le module d'encodage basé maillage.

15 Les images codées en mode maillage utilisent comme références pour leur prédiction temporelle les images codées en mode MPEG-4 située immédiatement avant ou immédiatement après le groupe d'images codées en mode maillage auquel elles appartiennent.

20 Le point clé de l'efficacité en compression de l'invention est que la compensation de mouvement basée maillage conduit à une prédiction temporelle très performante, pour un coût de codage associée très faible.

En effet, cette technique :

- Prend en compte différents types de mouvements dans les images
- Traite proprement les recouvrements et découvrements de zones dus aux mouvements d'objets.

25 La figure 1 donne une vue générale du principe du codeur.

Tout d'abord, les images entrantes sont aiguillées soit vers le module d'encodage MPEG, soit vers le module d'encodage base maillage, selon un mode de décision donné, par exemple selon un rythme prédéfini : 1 image sur N est codée en MPEG, les autres en mode interpolé maillage. On note N_k les numéros
30 des images codées en mode MPEG.

Toutes les autres images I_i , $N_k < i < N_{k+1}$, sont encodées par un encodeur basé sur un maillage, par exemple triangulaire, fonctionnant en mode interpolé dit mode B. le principe général de cet encodeur est le suivant :

1. Calcul des champs de mouvement avant et arrière entre les images N_k et N_{k+1} . Ces champs sont modélisés sous forme de maillages, triangulaires.
- 2a. Estimation des zones prédictibles : lors de l'interpolation, on estime le mouvement entre I_t et $I_{t'}$. Si l'application des vecteurs mouvements ne conduit à aucun retournement (ce qui signifie que l'on est en présence d'une zone prédictible), on réalise une compensation de mouvement en pondérant les vecteurs par un scalaire k ($0 \leq k \leq 1$) de manière à interpoler I_{t+k} avec $t+k$ appartenant à $]t, t[$. Trois approches sont envisageables pour la compensation de mouvement. Elles sont décrites par la suite ;
- 2b. Estimation des zones non prédictibles : Détection des zones d'occultation, non prédictibles, dans les images I_i à coder, à partir de la connaissance de ces champs de mouvement
3. Codage spécifique de ces zones d'occultation selon l'un des trois modes possibles suivants :
 - Prédiction avec l'une des images de référence (N_k , N_{k+1} , ou ces images compensées en mouvement avec leurs champs de mouvement) sans compensation de mouvement, puis codage de l'erreur de prédiction avec une technique basée maillage triangulaire
 - Prédiction avec l'une des images de référence (N_k , N_{k+1} , ou ces images compensées en mouvement avec leurs champs de mouvement) avec compensation de mouvement intra-image, puis codage de l'erreur de prédiction avec une technique basée maillage triangulaire
 - Codage intra-image basé avec une technique basée maillage

triangulaire.

4. Optionnellement, codage de type MPEG mode P de l'erreur résiduelle de prédiction ou de codage, limité à une zone autour de la zone d'occultation.

5 Comme mentionné plus haut, la compensation de mouvement peut être effectuée selon trois approches : avec une estimation avant, avec une estimation arrière ou avec une estimation avant et une estimation arrière.

1) avec une estimation avant :

10 Lors de l'interpolation, on estime le mouvement entre I_{t1} et I_{t2} . Si l'application des vecteurs mouvement ne conduit à aucun retournement (ce qui correspond aux zones prédictibles) alors on réalise une compensation de mouvement en pondérant les vecteurs par un scalaire $k = m/(t1+t2)$ ($0 < k < 1$) de manière à interpoler I_m avec m appartenant à $[t1, t2]$. On obtient l'image estimée E1.

15 2) avec une estimation arrière :

Lors de l'interpolation, on estime le mouvement entre I_{t2} et I_{t1} . Si l'application des vecteurs mouvements ne conduit à aucun ce qui correspond aux zones prédictibles) alors on réalise une compensation de mouvement en pondérant les vecteurs par un scalaire $k' = 1-m/(t1+t2)$ ($0 < k' < 1$) de manière à interpoler I_m avec m appartenant à $[t1, t2]$. On obtient l'image estimée E2.

3) les deux ensemble :

25 Lors de l'interpolation, on estime le mouvement I_{t1} et I_{t2} et entre I_{t2} et I_{t1} . Si l'application des vecteurs mouvements ne conduit à aucun retournement (ce qui correspond aux zones prédictibles), alors on réalise une compensation de mouvement en pondérant les vecteurs "avant" par un scalaire k ($0 < k < 1$) et l'on obtient E1. On fait de même avec les vecteurs "arrière" par un scalaire k' ($0 < k' < 1$) et on obtient E2. L'image estimée est alors $E = aE1 + (1-a) E2$ avec $0 < a < 1$.

30 On prend à chaque fois moyennant 2 bits la meilleure des solutions après

calcul du PSNR entre la solution estimée et l'image source associée.

1. Calcul des champs de mouvement avant et arrière entre les images

N_k et N_{k+1}

Des champs de mouvement avant et arrière entre les images N_k et N_{k+1} sont
5 calculés, sous la forme de maillages hiérarchiques, par exemple triangulaires, T_k^a
et T_{k+1}^a , comme indiqué figure 2.

De tels maillages sont obtenus par division de certaines mailles, par
exemple, les mailles triangulaires sont divisées en 4 sous-triangles, en fonction
d'un certain critère au cours du processus d'estimation du mouvement. À chaque
10 niveau de la hiérarchie, les décisions de division ou non sont prises pour chaque
maille. Une fois ces divisions décidées, les mailles adjacentes des mailles divisées
sont alors divisées de sorte à conserver une structure de maillage conforme. Le
maillage initial, avant division (sommet de la hiérarchie), peut être quelconque.

Dans l'exemple de la figure 2, l'estimateur de mouvement décide de diviser
15 les triangles 3 et 8. ceci entraîne la division des triangles 2, 4, 7 et 9. le processus
est itéré jusqu'à un niveau prédéfini de hiérarchie.

Dans le cas de maillages triangulaire, l'expression du champ de
mouvement défini par un maillage triangulaire T est donné sur chaque triangle e
par :

$$20 \quad \vec{d}(\vec{p}, D) = \sum_{n \in \text{ner}(e)} \Psi_n^e(\vec{p}) \vec{d}_n, \quad \forall \vec{p}(x, y) \in e$$

où :

- e dénote l'élément triangulaire de T contenant le point courant $\vec{p}$ de
coordonnées x et y,
- $\text{ner}(e)$ dénote l'ensemble de ses trois nœuds ou sommets, numérotés i, j,
25 k de positions $\vec{p}_i, \vec{p}_j$ et $\vec{p}_k$,
- Ψ_l ($l = i, j, k$) représente les coordonnées barycentriques du point $\vec{p}(x, y)$
dans l'élément triangulaire $e_{i,j,k}$ avec :

$$\begin{cases} \Psi_l^e(x, y) = \alpha_l^e + \beta_l^e x + \gamma_l^e y, & \alpha_l^e, \beta_l^e, \gamma_l^e \in \mathbb{R} \\ \sum_{l=i,j,k} \Psi_l^e(x, y) = 1 & \text{si } \vec{p}(x, y) \in e_{i,j,k} \\ \Psi_l^e(x, y) = 0 & \text{sinon} \end{cases}$$

Un tel modèle définit un champ partout continu. De plus, il permet un contrôle fin de la précision de représentation, caractéristique essentielle pour la compression.

À chaque niveau de la hiérarchie de maillages, les vecteurs nodaux de mouvement sont calculés de sorte à minimiser une erreur de prédiction. Différents estimateurs de mouvement basés maillages peuvent être utilisés, par exemple celui décrit dans le brevet FR n° 98 11227, ou FR n° 99 15568.

Le point important est que le maillage final résulte d'un processus hiérarchique à partir d'un maillage initial par divisions. Ce caractère hiérarchique est en effet mis à profit pour le codage différentiel des vecteurs nodaux de mouvement entre un nœud et ses nœuds parents (les extrémités de l'arc sur lequel il a été inséré). La structure du maillage est recalculée au décodeur à partir de la connaissance du maillage initial, et des indicateurs de division de mailles.

Ainsi, en fin de processus, 2 maillages de mouvement sont obtenus pour chaque groupe d'images comprises entre les images N_k et N_{k+1} , utilisés pour reconstruire l'ensemble des images du groupe.

2. Détection des zones d'occultation

À partir de ces 2 maillages, les zones d'occultation, c'est-à-dire non prédictibles dans l'image N_k à partir de l'image N_{k+1} ou vice-versa, du fait du recouvrement ou du découvrément d'objets, sont détectées.

Ces zones sont simplement définies par les triangles se recouvrant, une fois déplacés par leurs vecteurs nodaux.

La figure illustre la détection d'occultation base sur le recouvrement de triangles après déplacement.

Le codeur peut poursuivre l'estimation de mouvement en désactivant les triangles des zones d'occultation, de manière à obtenir des vecteurs déplacement moins biaisés.

Ceci est cependant strictement interne à la stratégie du codeur, et au final, c'est 2 maillages de mouvement complets T_k^0 et T_{k+1}^1 qui sont codés et insérés dans le train binaire. Le décodeur est alors capable de retrouver les zones d'occultation

à partir de ces 2 maillages.

Ces zones d'occultation sont définies sur les images N_k et N_{k+1} et une fois détectés, les triangles leur appartenant sont étiquetés en conséquence, au codeur comme au décodeur.

5 Cependant, le codeur a besoin de connaître ces zones sur les images N_{k+1} à $N_{k+1}-1$. Celles-ci sont simplement obtenues par projection des maillages T_k^b et T_{k+1}^f sur l'image à coder, par application des vecteurs mouvement nodaux, renormalisés pour tenir compte de la distance temporelle entre l'image courante et l'image de référence N_k ou N_{k+1} .

10 3. Codage des zones d'occultation :

Pour chaque zone d'occultation, l'image de référence pour une éventuelle prédiction est sélectionnée parmi I_{Nk} , $I_{N(k+1)}$, mais aussi I_1 , qui est l'image obtenue à l'instant 1 par compensation de mouvement avec le maillage T_k^b ou T_{k+1}^f à un niveau où il n'y a pas encore de recouvrement de maille.

15 Plus précisément, le choix entre I_{Nk} et $I_{N(k+1)}$ dépend simplement du maillage T_k^b ou T_{k+1}^f ayant engendré la zone courante d'occultation. Ensuite, ce maillage est utilisé pour prédire l'image et donner I_1 . Le choix entre I_1 et I_{Nk} ou $I_{N(k+1)}$ se fait sur critère d'erreur de prédiction : l'image donnant l'erreur la plus faible est retenue. Ainsi, il suffit d'insérer 1 bit dans le flux binaire, par zone, pour

20 coder le choix de la prédiction retenue.

Notons I_r l'image de référence sélectionnée.

Le reste du codage de ces zones consiste en 2 étapes :

- Une étape de prédiction
 - Une étape de codage de l'erreur de prédiction ou de la texture originale
- 25 en cas de mauvaise prédiction

3.1 Prédiction résiduelle de la texture des zones d'occlusion

3 modes peuvent être utilisés, de manière exclusive. La décision est prise sur critère de moindre erreur.

Mode 1:

30 Les valeurs Y, U et V d'un pixel de la zone sont simplement celle du pixel

de même localisation de l'image de référence I_r . Soit $\hat{I}_1$ l'image résultante. On code alors l'erreur de prédiction entre $\hat{I}_1$ et I_l .

Mode 2:

5 Une estimation de mouvement est alors réalisée entre I_r (l'image à coder) et $\hat{I}_1$ (le résultat de la prédiction du mode 1) sur la zone d'occultation. Le maillage résultant, issu du dernier niveau du maillage T_r , $r=k$ ou $k+1$, avant les recouvrements de mailles, est alors codé ainsi que ses mouvements nodaux. Enfin, l'erreur résiduelle de prédiction est codée selon une procédure définie plus loin.

Mode 3:

10 On n'effectue aucune prédiction et on code les valeurs originales des pixels de la zone.

4. Codage de la texture ou de l'erreur de prédiction sur les zones d'occultation.

15 La texture originale et l'erreur de prédiction subissent le même codage, dont le principe est le suivant :

On suppose qu'un maillage triangulaire initial a pu être défini à partir du maillage de mouvement T_r retenu pour la prédiction de la zone à coder. La façon de dériver ce maillage initial sera décrit plus loin.

La texture est alors approximée sur chaque maille en fonction d'un choix:

- 20
- Les mailles riches en hautes fréquences sont codées à base de transformée en cosinus discrète, dite DCT
 - Les mailles plus lisses sont codées par un modèle d'éléments finis, affine.

25 Là encore, on va tirer parti d'une approche hiérarchique pour diminuer le coût de codage de la représentation par maillage.

L'approche retenue permet de conserver le faible coût de codage associé à une hiérarchie régulière de mailles tout en permettant l'adaptation locale au contenu des images permise par la décomposition irrégulière de mailles.

30 À partir du maillage initial grossier de la zone, les mailles sont subdivisées en 4 sous-mailles triangulaires jusqu'à un niveau donné.

Sur le dernier niveau, une permutation optionnelle des diagonales des quadrilatères engendrés par 2 triangles adjacents peut être mise en œuvre, si cela induit une baisse de l'erreur d'approximation.

4.1 Initialisation du maillage de texture sur les zones d'occultation

5 Ce maillage est simplement donné par le dernier niveau de T_r (maillage résultant du déplacement de T_k ou T_{k+1} selon le sens retenu) avant apparition des retournements sur la zone considérée. Ainsi, on a un maillage de texture qui s'insère de manière naturelle dans le maillage mouvement, puisque extrait de ce dernier.

4.2 Représentations utilisées pour la texture sur les triangles

10 2 représentations sont combinées : l'interpolation affine et la DCT 4triangulaire.

Interpolation affine

15 Les nœuds du maillage triangulaire portent l'information photométrique (couleur, erreur) et l'interpolation pour les points à l'intérieur du triangle est effectuée par un élément fini de Lagrange, également appelé interpolation affine.

La valeur $v(\vec{p})$ du point $\vec{p}(x,y)$ intérieur au triangle $e_{i,j,k}$ défini par les 3 nœuds $\vec{p}_l, l = i, j, k$ est donnée par l'équation suivante :

$$v(\vec{p}) = \sum_{l=i,j,k} \Psi_l(x,y)(\vec{p})$$

où Ψ_l ($l = i, j, k$) représente les coordonnées barycentriques du point.

20 $v(\vec{p})$ peut être l'une des composantes photométriques Y, U ou V du point, ou encore l'erreur de prédiction pour ces composantes.

Plusieurs méthodes peuvent être employées pour le calcul des valeurs nodales, en particulier la méthode des moindres carrés.

Transformation en cosinus discrète (DCT) sur les triangles

25 Le principe de la méthode consiste à transformer un triangle quelconque en un triangle de référence isocèle rectangle. Le contenu de ce triangle est alors symétrisé par rapport à l'hypoténuse pour donner une matrice carrée symétrique (figure 4).

Une DCT classique (carrée) est alors appliquée à cette matrice. On peut

montrer que la matrice transformée est aussi symétrique. Seuls les coefficients de son triangle inférieur sont alors quantifiés puis codés statistiquement (codage entropique).

La figure 4 décrit les différentes étapes du processus : sélection du triangle T, transformation affine de celui-ci en un triangle isocèle rectangle T'. Du fait de la transformation affine, les pixels du triangle ne sont plus situés sur une grille orthogonale régulière, et il convient de ré-échantillonner les valeurs photométriques de l'intérieur du triangle de référence. On utilise pour cela un processus analogue à celui d'une compensation de mouvement dans l'image (en l'occurrence la transformation affine), à l'aide d'un interpolateur, par exemple bilinéaire.

La transformation affine F et son inverse F^{-1} sont définies par les équations suivantes :

$$F : \begin{cases} X = \frac{(x_2 - x_1)(y_1 - y) + (y_1 - y_2)(x_1 - x)}{(x_2 - x_1) + (y_2 - y_1) + (y_1 - y_2)(x_2 - x_1)} N \\ Y = \frac{(x_1 - x_2)(y_1 - y) + (y_2 - y_1)(x_1 - x)}{(x_2 - x_1) + (y_2 - y_1) + (y_1 - y_2)(x_2 - x_1)} N \end{cases}$$

$$F^{-1} : \begin{cases} x = x_1 + (x_2 - x_1) \frac{X}{N} + (x_3 - x_1) \frac{Y}{N} \\ y = y_1 + (y_2 - y_1) \frac{X}{N} + (y_3 - y_1) \frac{Y}{N} \end{cases}$$

Les valeurs photométriques $M(i, j)$ du triangle T' (donc du bloc M, symétrisé de T') sont obtenues par transformation inverse F^{-1} puis interpolation Γ de la texture de l'image à coder :

$$M(i_T, j_X) = I_\Gamma(P_{(x,y)}) = I_\Gamma(F^{-1}(Q_{(x,y)}))$$

où :

le coefficient $M(i_T, j_X)$ est la valeur au point $Q(X, Y)$ dont le transformé $P(x, y)$ est $F^{-1}(Q)$

I_Γ dénote l'interpolateur utilisé pour calculer la valeur de l'image au point, de coordonnées potentiellement non entières.

La reconstruction $\hat{\Gamma}$ de la texture Γ est donnée par :

$$\hat{\Gamma}(P_{(x,y)}) = I_\Gamma(Q_{(x,y)}) = I_\Gamma(F(P_{(x,y)}))$$

où I_p dénote la texture interpolée à partir des valeurs du bloc M' , version quantifiée de M .

Cette technique ne peut être appliquée qu'aux seuls triangles d'aire non nulle. Mais de tels triangles ne nécessitent pas de codage de texture par définition.

5 Contrairement à la SADCT (DCT adaptée à une forme), cette transformation ne garantit pas la reconstruction parfaite après transformation inverse, même en l'absence de quantification.

Afin de réduire l'erreur de reconstruction, un facteur d'échelle α est introduit pour le calcul du bloc M_i (de taille $N_i \times N_i$) pour chaque triangle i :

$$10 \quad N_i = E\left(\alpha \sqrt{2A_i}\right)$$

où :

- E est la partie entière par excès,
- A_i est l'aire du triangle i .

15 En fait, $\alpha = 1$ réalise un compromis intéressant, qui est plus efficace pour les triangles proche d'un isocèle. Le cas $\alpha < 1$ est utilisé conjointement au pas de quantification pour comprimer le volume d'information.

Une fois le bloc M_i défini pour chaque triangle, une transformation DCT classique leur est appliquée, et les coefficients transformés sont quantifiés selon plusieurs procédés possibles, par exemple une quantification scalaire uniforme, ou 20 encore une quantification incrémentale avec la fréquence du coefficient transformée. L'emploi des matrices de quantification MPEG ou JPEG bien connues est également possible.

L'expression de la DCT est donnée par :

$$25 \quad F(u, v) = \frac{1}{\sqrt{NM}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \lambda(u) \lambda(v) \cos \frac{(2i+1)\pi}{2N} \cos \frac{(2j+1)\pi}{2M} f(i, j)$$

$$\lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } \xi = 0 \\ 1 & \text{sinon} \end{cases}$$

On a la relation $F(u, v) = F(v, u)$ car :

$$f(i, j) = f(j, i) \quad (\forall u, v, i, j = 0, \dots, N-1) \text{ par définition.}$$

Par conséquent, on peut se contenter de ne calculer que les coefficients de la partie inférieure de la matrice transformée.

4.3 codages global de texture

Comme indiqué précédemment, on utilise un maillage hiérarchique uniforme obtenu par division de chaque triangle d'un niveau donné de la hiérarchie en 4 sous-triangles, par insertion de nœuds au milieu des arcs. Le processus est répété itérativement jusqu'à un niveau maximal. Cette hiérarchie de triangles est également représentée et gérée par le codeur sous forme d'un arbre quaternaire (figure 5). Notons que seuls les triangles inclus dans la zone à coder sont pris en compte. Le procédé de construction du maillage initial de base garanti que tout triangle de la hiérarchie de maillage appartient à la zone à coder.

Le processus de codage par maillage d'une zone d'occultation peut être résumé de la manière suivante :

- ① un maillage hiérarchique imbriqué est défini sur la zone à coder, par création d'un maillage initial régulier puis sub-division itérative des triangles en 4 sous-triangles par insertion de nouveaux nœuds au milieu des arcs. Les valeurs aux nœuds sont calculées pour minimiser l'erreur d'approximation de la zone par le maillage ;
- ② les valeurs des pixels sont approchées par une interpolation affine sur le triangle les contenant à partir de ses valeurs aux nœuds.

Pour chaque triangle de la hiérarchie, on évalue alors l'erreur d'approximation E puis on décide des différents mode de représentation et codage en fonction de 2 seuils : σ_1 et σ_2 :

- ① si $E < \sigma_1$, l'interpolation affine est suffisante sur le triangle ;
 - ② si $\sigma_1 < E < \sigma_2$, il faut utiliser une décomposition plus fine du triangle pour obtenir une bonne approximation, toujours par interpolation affine.
 - ③ si $E > \sigma_2$, le triangle est texturé et on code l'erreur d'interpolation affine en utilisant la DCT.
- Enfin, sur le maillage le plus fin, on teste la diminution d'erreur apportée

par la permutation de diagonale des quadrilatères formés par 2 triangles adjacents. En cas de résultat positif, on valide cette permutation.

Selon les modes de codage choisis pour les différents triangles, on code les différentes informations de la manière suivantes

5 Les valeurs nodales YUV sont d'abord prédites à partir des valeurs des nœuds parents (extrémités de l'arc où on a inséré le nœud courant). La différence entre la valeur du nœud et sa valeur prédite est alors quantifiée.

Enfin, la structure de l'arbre quaternaire (incluant les indicateurs de division ou non des triangles), les indicateurs de permutation de diagonales, les
10 valeurs différentielles nodales de YUV et les coefficients DCT quantifiées sont codés par un codeur arithmétique et insérés dans le train binaire.

5. Résumé des informations codées dans le flux binaire des trames codées par maillage

Chaque groupe de trames encodées en mode maillage entre N_k+1 et $N_{k+1}-1$
15 (où N_k , N_{k+1} sont respectivement la trame précédente et la trame suivante encodée en mode MPEG) est représenté comme un tout dans le flux binaire.

L'information véhiculée comprend, sous forme codée :

Une entête pour l'ensemble du groupe de trames, incluant entre autres le nombre réel de trames encodées.

20 Les maillages de mouvement (structure et vecteurs déplacement des nœuds) T_k^b et T_{k+1}^f .

La texture d'erreur de prédiction ou originale, pour chaque image du groupe.

6. Structure globale du train binaire

25 Le train binaire global consiste en une succession de trames encodées en mode MPEG, et de groupes de trames encodés en mode maillage, comme indiqué figure 8.

L'entête globale du train binaire représentatif de la séquence codée contient entre autres l'indication de codage hybride.

30 La partie du train binaire correspondant à un groupe de trames codées en

mode maillage débute par une entête indiquant entre autres le nombre de trame effectivement codées, éventuellement nul.

Les différents flux de données (trains binaires) correspondant respectivement à l'entête globale de la séquence codée, aux images encodées MPEG et aux groupes d'i images encodées en mode interpolé par maillage peuvent être envoyés sur différentes voies indépendantes si besoin. En particulier, le procédé de codage permet le décodage hiérarchique (ou scalable) de la séquence, c'est-à-dire un décodage n'utilisant qu'une partie du débit total.

7. Processus de décodage

La figure 9 donne une vue générale du principe du décodage.

Tout d'abord, le décodage de l'entête permet d'activer le décodage hybride.

Ensuite, le décodeur reconnaît pour chaque partie du train binaire correspondant à une entité autonome s'il s'agit d'une trame encodée MPEG-4 ou d'un groupe de trames encodées par maillage.

Les trames MPEG-4 sont fournies au module de décodage MPEG-4, et les groupes de trames codées en mode maillage sont fournies au module de décodage par maillage.

7.1 Décodage basé maillage

Tout d'abord, les maillages de mouvement $T_{k,x}^b$ et $T_{(k+1)}^f$ pour le groupe d'images I_i , $N_k < i < N_{k+1}$ sont décodés.

Ensuite, les zones d'occultation pour ces images sont retrouvées selon le même procédé qu'au codeur.

Les pixels en dehors des zones d'occultation sont simplement interpolés à partir des images I_{N_k} et $I_{N_{k+1}}$ et des champs de mouvement $T_{k,x}^b$ et T_{k+1}^f .

Le maillage de texture le plus grossier (sommet de la hiérarchie) est retrouvé pour chaque zone d'occultation selon un procédé identique à celui du codeur.

L'information associée au maillage hiérarchique correspondant (indicateur de division de triangles, décisions d'interpolation affine ou de codage DCT, valeur YUV nodale différentielles et coefficients DCT quantifiés) est alors décodée et les

WO 02/060184

21

PCT/FR02/00322

valeurs YUY des pixels de ces zones reconstruits.

REVENDICATIONS

1. Procédé de codage d'images, caractérisé en ce qu'il met en œuvre sélectivement au moins deux modes de codage d'images, optimisant chacun la compression d'au moins une image d'une séquence vidéo en fonction de critères d'optimisation différents.
2. Procédé de codage selon la revendication 1, caractérisé en ce qu'une information sur le choix d'un desdits modes de codage est connue d'un décodeur selon au moins une des techniques appartenant au groupe comprenant :
- choix prédéfini, connu au codage et au décodage ;
 - information représentative du choix incluse dans un flux de données comprenant au moins certaines des données d'images codées ;
 - information représentative du choix incluse dans un flux de données indépendant des données d'images codées ;
 - détermination du choix de façon intrinsèque, par le décodeur.
3. Procédé de codage selon l'une quelconque des revendications 1 et 2, caractérisé en ce qu'il comprend une étape de sélection d'un mode de codage à appliquer à ladite image, parmi au moins :
- un premier codage optimisant sensiblement une représentation photométrique d'une image ;
 - un second codage optimisant sensiblement une représentation du mouvement entre au moins deux images.
4. Procédé de codage selon la revendication 3, caractérisé en ce que ledit second codage tient compte d'au moins une image précédente et/ou au moins une image suivante codée à l'aide dudit premier codage.
5. Procédé de codage selon la revendication 4, caractérisé en ce que ledit second codage tient compte d'un champ de vecteurs mouvement calculé à partir de l'image immédiatement précédente codée à l'aide dudit premier codage et/ou d'un champ de vecteurs mouvement calculé à partir de l'image

immédiatement suivante codée à l'aide dudit premier codage.

6. Procédé de codage selon la revendication 5, caractérisé en ce que ledit champ de vecteurs mouvement est appliqué à un maillage.

7. Procédé de codage selon l'une quelconque des revendications 5 et 6, caractérisé en ce que lesdits champs de vecteurs mouvement sont utilisés pour déterminer un champ de vecteurs mouvement déduit, associé à une image codée à l'aide dudit second codage.

8. Procédé de codage selon l'une quelconque des revendications 5 à 7, caractérisé en ce que ledit second codage met en œuvre une estimation de mouvement avant, entre une image I_{t1} et une image suivante I_{t2} , et une étape de compensation de mouvement dans laquelle les vecteurs mouvement obtenus lors de ladite estimation de mouvement et ne conduisant à aucun retournement sont pondérés par un scalaire $k = m/(t1 + t2)$, $0 < k < 1$, de façon à interpoler au moins une image $Im1$, m appartenant à $[t1, t2]$.

9. Procédé de codage selon l'une quelconque des revendications 5 à 7, caractérisé en ce que ledit second codage met en œuvre une estimation de mouvement arrière, entre une image I_{t2} et une image précédente I_{t1} , et une étape de compensation de mouvement dans laquelle les vecteurs mouvement obtenus lors de ladite estimation de mouvement et ne conduisant à aucun retournement sont pondérés par un scalaire $k' = 1 - m/(t1 + t2)$, $0 \leq k' \leq 1$, de façon à interpoler au moins une image $Im2$, m appartenant à $[t1, t2]$.

10. Procédé de codage selon les revendications 8 et 9, caractérisé en ce qu'il met en œuvre une estimation arrière et une estimation avant, de façon à obtenir une image estimée Im valant $Im = aE1 + (1-a)E2$ avec $0 \leq a \leq 1$.

11. Procédé de codage selon l'une quelconque des revendications 3 à 10, caractérisé en ce que ladite étape de sélection repose sur la mise en œuvre d'un sous-échantillonnage de facteur fixé N , une image sur N étant codée à l'aide dudit premier codage.

12. Procédé de codage selon la revendication 11, caractérisé en ce que N est supérieur à 2.

13. Procédé de codage selon l'une quelconque des revendications 11 et 12, caractérisé en ce que N est variable.
14. Procédé de codage selon l'une quelconque des revendications 3 à 11, caractérisé en ce que ledit premier codage met en œuvre une transformation sur des blocs d'images et une prédiction temporelle par blocs.
15. Procédé de codage selon la revendication 14, caractérisé en ce que ledit premier codage est un codage MPEG-4 ou H26L.
16. Procédé de codage selon la revendication 115, caractérisé en ce que les images délivrées par ledit codage MPEG-4 ou H26L comprennent des images de type I (intra) et/ou de type P (prédicatif).
17. Procédé de codage selon l'une quelconque des revendications 3 à 16, caractérisé en ce que ledit second codage repose sur la mise en œuvre d'un maillage hiérarchique de M niveaux, M étant supérieur ou égal à 1.
18. Procédé de codage selon la revendication 17, caractérisé en ce que ledit maillage est triangulaire.
19. Procédé de codage selon l'une quelconque des revendications 17 et 18, caractérisé en ce qu'il comprend une étape de gestion des zones d'occlusions.
20. Procédé de codage selon l'une quelconque des revendications 1 à 19, caractérisé en ce qu'il produit au moins deux flux de données, pouvant être transmis sur des voies de transmission indépendantes.
21. Procédé de codage selon la revendication 20, caractérisé en ce que lesdits flux de données appartiennent au groupe comprenant :
- une entête globale ;
 - des données d'image codées selon ledit premier codage ;
 - des données d'image codées selon ledit second codage.
22. Procédé de décodage d'un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 21.
23. Dispositif de codage d'un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 21.
24. Dispositif de décodage d'un signal d'images codé à l'aide du

procédé de codage de l'une quelconque des revendications 1 à 21.

25. Dispositif de décodage selon la revendication 24, caractérisé en ce qu'il comprend des moyens de détermination d'au moins une partie d'un champ de vecteurs et/ou d'au moins une partie des zones d'occlusion, similaires à ceux mis en œuvre lors du codage.

26. Dispositif de stockage d'au moins un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 21.

27. Système de codage, de transmission et/ou de décodage d'un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 17.

28. Système selon la revendication 27, caractérisé en ce qu'une information sur le choix d'un desdits modes de codage est connue d'un décodeur selon au moins une des techniques appartenant au groupe comprenant :

- choix prédéfini, connu au codage et au décodage ;
- information représentative du choix incluse dans un flux de données comprenant au moins certaines des données d'images codées ;
- information représentative du choix incluse dans un flux de données indépendant des données d'images codées ;
- détermination du choix de façon intrinsèque, par le décodeur.

29. Produit programme d'ordinateur pour le codage et/ou le décodage d'un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 21.

30. Support de données portant un programme d'ordinateur pour le codage et/ou le décodage d'un signal d'images codé à l'aide du procédé de codage de l'une quelconque des revendications 1 à 21.31. Signal de données d'images, caractérisé en ce qu'il comprend des données codées selon le procédé de l'une quelconque des revendications 1 à 21.

32. Signal selon la revendication 31, caractérisé en ce qu'au moins un

indicateur indiquant si le procédé selon l'une quelconque des revendications 1 à 21 est ou non activé.

5 **33.** Signal selon l'une quelconque des revendications 31 et 32, caractérisé en ce qu'il comprend une donnée précisant la structure des trames, en début de séquence vidéo *et/ou* dans chaque trame de signal.

34. Signal selon l'une quelconque des revendications 31 à 33, caractérisé en ce qu'une séquence codée à l'aide dudit second codage débute par une entête précisant le nombre de trames codées selon ce second codage.

10 **35.** Signal selon l'une quelconque des revendications 31 à 34, caractérisé en ce qu'il comprend au moins deux flux de données, pouvant être transmis sur des voies de transmission indépendantes.

36. Signal selon la revendication 35, caractérisé en ce que lesdits flux de données appartiennent au groupe comprenant :

- 15 - une entête globale ;
 - des données d'image codées selon ledit premier codage ;
 - des données d'image codées selon ledit second codage.

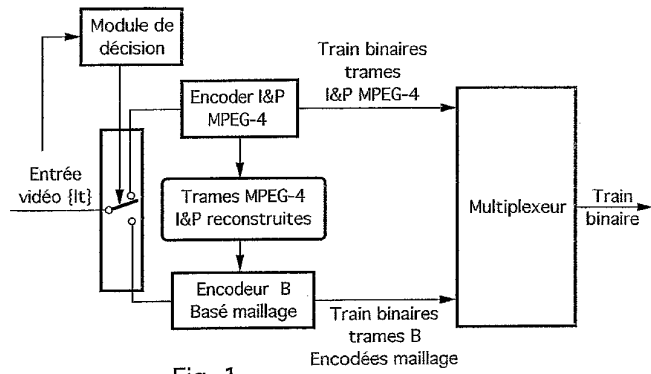
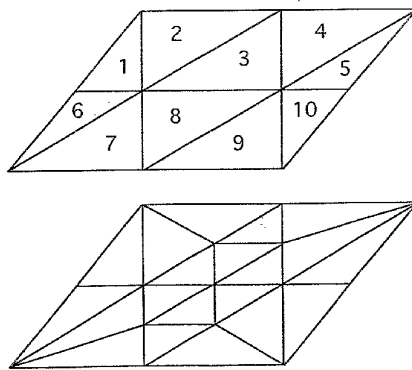
37. Application du procédé de codage selon l'une quelconque des revendications 1 à 21 à au moins un des domaines appartenant au groupe comprenant :

- 20 - la télévision numérique ;
 - la vidéo temps réel sur réseau IP ;
 - la vidéo temps réel sur réseau vers les mobiles ;
 - le stockage de données d'images.

WO 02/060184

PCT/FR02/00322

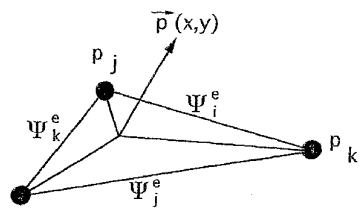
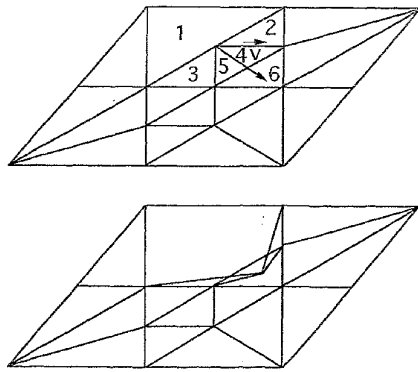
1/5

Fig. 1Fig. 2

WO 02/060184

PCT/FR02/00322

2/5

Fig. 3Fig. 4

WO 02/060184

PCT/FR02/00322

3/5

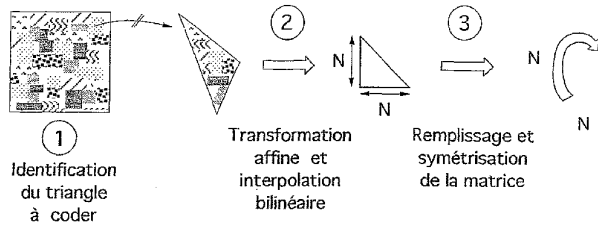


Fig. 5

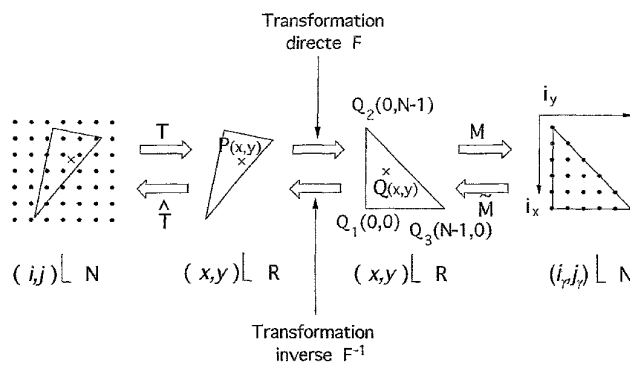
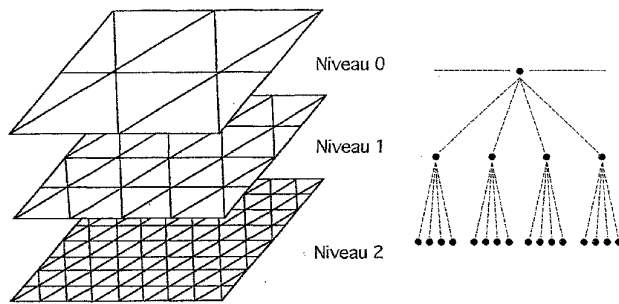
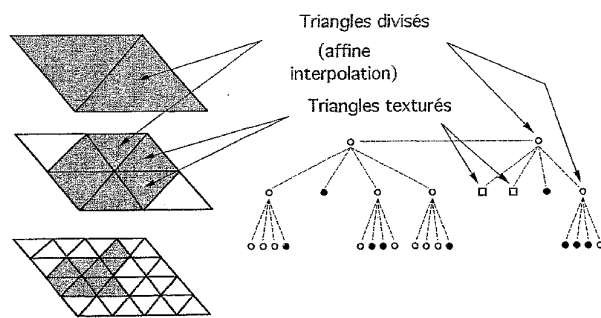


Fig. 6

WO 02/060184

4/5

PCT/FR02/00322

Fig. 7Fig. 8

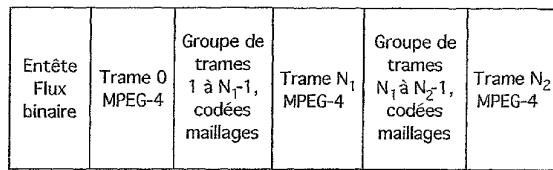


Fig. 9

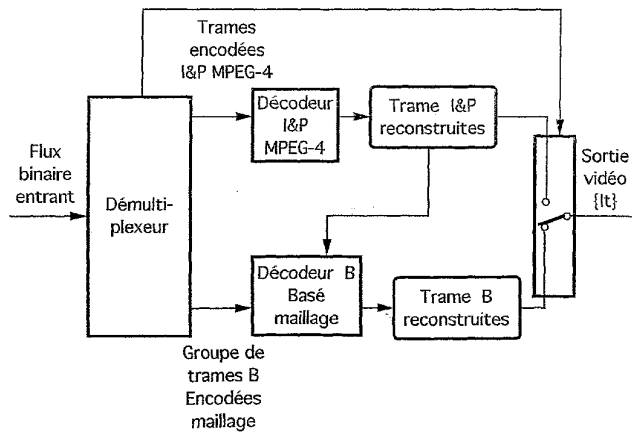


Fig. 10

【国際調査報告】

INTERNATIONAL SEARCH REPORT		Int. national Application No PCT/FR 02/00322
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 H04N7/46 H04N7/36		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, COMPENDEX, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 771 117 A (DAE WOO ELECTRONICS CO LTD) 2 May 1997 (1997-05-02) page 4, line 5 - line 28; figure 1 page 4, line 43 -page 6, line 30; figures 3,4 --- -/--	1,3-5,7, 11-14, 20-24, 26,27, 29-37 2,6, 8-10, 15-19, 25,28
A		
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention ** document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *** document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *S* document member of the same patent family		
Date of the actual completion of the international search 12 June 2002		Date of mailing of the international search report 26/06/2002
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Georgiou, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/FR 02/00322
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 978 030 A (JUNG HAE-MOOK ET AL) 2 November 1999 (1999-11-02) column 6, line 19 -column 7, line 44; figure 1 column 9, line 45 -column 10, line 49; figure 3	1,3-5,7, 11-14, 20-24, 26,27, 29-37
A		2,6, 8-10, 15-19, 25,28
Y	TUDOR P N: "MPEG-2 VIDEO COMPRESSION TUTORIAL" IEEE COLLOQUIUM ON MPEG WHAT IT IS AND WHAT IT ISN'T, IEEE, LONDON, GB, 1995, pages 2-1-2-8, XP000562433 * alinéa 'Picture types! *	2,28
Y	"MPEG-4 OVERVIEW - (MAUI VERSION)" ISO/IEC JTC1/SC29/WG11 N3156, XX, XX, December 1999 (1999-12), pages 1-57, XP001009690 * alinéa '2.5.8.1 Resynchronization! *	2,28
Y	WO 97 39577 A (SOMERVILLE STUART ;KNEE MICHAEL JAMES (GB); SNELL & WILCOX LTD (GB) 23 October 1997 (1997-10-23) page 3, line 27 -page 5, line 7	2,28
A	NEWCZLOWSKI J ET AL: "MOTION COMPENSATED VIDEO SEQUENCE INTERPOLATION USING DIGITAL IMAGEWARPING" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP). I. IMAGE AND MULTIDIMENSIONAL SIGNAL PROCESSING. ADELAIDE, APR. 19 - 22, 1994, NEW YORK, IEEE, US, vol. 5 CONF. 19, 19 April 1994 (1994-04-19), pages V-205-V-208, XP000522474 ISBN: 0-7803-1776-9 paragraphs '0III!', 'VI.A!	1-37
	-/--	

INTERNATIONAL SEARCH REPORT		International Application No. PCT/FR 02/00322
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LECHAT P ET AL: "COMBINED MESH BASED IMAGE REPRESENTATION AND MOTION ESTIMATION, APPLICATION TO VIDEO CODING" PROCEEDINGS OF THE 1998 INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, ICIP '98, CHICAGO, IL, OCT. 4 - 7, 1998, INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, LOS ALAMITOS, CA: IEEE COMPUTER SOC. US, vol. 2 CONF. 5, 4 October 1998 (1998-10-04), pages 909-913, XP000827604 ISBN: 0-8186-8822-X paragraphs '0002!', '0004!', '0005!; figure 5	1-37
A	US 4 383 272 A (NETRAVALI ARUN N ET AL) 10 May 1983 (1983-05-10) column 3, line 24 -column 4, line 13; figure 2	1-37
A	DEPOMMIER R ET AL: "MOTION ESTIMATION WITH DETECTION OF OCCLUSION AREAS" MULTIDIMENSIONAL SIGNAL PROCESSING. SAN FRANCISCO, MAR. 23 - 26, 1992, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), NEW YORK, IEEE, US, vol. 3 CONF. 17, 23 March 1992 (1992-03-23), pages 269-272, XP000378925 ISBN: 0-7803-0532-9 paragraph '0001!; figure 1	1-37
A	LAURENT N: "Hierarchical mesh-based global motion estimation, including occlusion areas detection" INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP 2000);VANCOUVER, BC, CANADA SEP 10-13 2000, vol. 3, 2000, pages 620-623, XP002201979 IEEE Int Conf Image Process;IEEE International Conference on Image Processing 2000 the whole document	1-37
A	GB 2 279 531 A (SONY UK LTD) 4 January 1995 (1995-01-04) page 9, line 30 -page 10, line 20; figure 6	1-37

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/FR 02/00322

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0771117	A	02-05-1997	KR 209793 B1 15-07-1999 CN 1152844 A 25-06-1997 EP 0771117 A2 02-05-1997 JP 9191460 A 22-07-1997 US 5668600 A 16-09-1997
US 5978030	A	02-11-1999	KR 181034 B1 01-05-1999 AT 207279 T 15-11-2001 AU 711311 B2 07-10-1999 AU 2420695 A 08-10-1996 BR 9510541 A 09-06-1998 CA 2215386 A1 26-09-1996 CZ 9702902 A3 18-02-1998 DE 69523340 D1 22-11-2001 DE 69523340 T2 02-05-2002 EP 0815690 A1 07-01-1998 FI 973715 A 17-11-1997 JP 11506576 T 08-06-1999 WO 9629828 A1 26-09-1996 NO 974209 A 31-10-1997 PL 322302 A1 19-01-1998
WO 9739577	A	23-10-1997	AU 2518497 A 07-11-1997 EP 0893025 A1 27-01-1999 WO 9739577 A1 23-10-1997
US 4383272	A	10-05-1983	NONE
GB 2279531	A	04-01-1995	JP 8009339 A 12-01-1996

RAPPORT DE RECHERCHE INTERNATIONALE		De n° internationale No PCT/FR 02/00322
A. CLASSEMENT DE L'OBJET DE LA DEMANDE CIB 7 H04N7/46 H04N7/36		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 7 H04N		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Bases de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés) EPO-Internal, COMPENDEX, INSPEC		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	EP 0 771 117 A (DAE WOO ELECTRONICS CO LTD) 2 mai 1997 (1997-05-02) page 4, ligne 5 - ligne 28; figure 1 page 4, ligne 43 -page 6, ligne 30; figures 3,4	1,3-5,7, 11-14, 20-24, 26,27, 29-37
A	---	2,6, 8-10, 15-19, 25,28
	-/-	
☒ Voir la suite du cadre C pour la fin de la liste des documents ☒ Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités: *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent *E* document antérieur, mais publié à la date de dépôt international ou après cette date *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison épistémologique (telle qu'indiquée) *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée ** document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre la portée ou la théorie constituant la base de l'invention *X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément ** document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier *A* document qui fait partie de la même famille de brevets		
Date à laquelle la recherche internationale a été effectivement achevée 12 juin 2002		Date d'expédition du présent rapport de recherche internationale 26/06/2002
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel: (+31-70) 340-2440, Tx: 31 851 epo nl, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Georgiou, G

Formulaire PCT/ISA/210 (deuxième feuille) (juillet 1992)

RAPPORT DE RECHERCHE INTERNATIONALE

De n° de internationale No
PCT/FR 02/00322

C(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	US 5 978 030 A (JUNG HAE-MOOK ET AL) 2 novembre 1999 (1999-11-02) colonne 6, ligne 19 -colonne 7, ligne 44; figure 1 colonne 9, ligne 45 -colonne 10, ligne 49; figure 3	1,3-5,7, 11-14, 20-24, 26,27, 29-37
A		2,6, 8-10, 15-19, 25,28
Y	TUDOR P N: "MPEG-2 VIDEO COMPRESSION TUTORIAL" IEE COLLOQUIUM ON MPEG WHAT IT IS AND WHAT IT ISN'T, IEE, LONDON, GB, 1995, pages 2-1-2-8, XP000562433 * alinéa 'Picture types! *	2,28
Y	"MPEG-4 OVERVIEW - (MAUI VERSION)" ISO/IEC JTC1/SC29/WG11 N3156, XX, XX, décembre 1999 (1999-12), pages 1-57, XP001009690 * alinéa '2.5.8.1 Resynchronization! *	2,28
Y	WO 97 39577 A (SOMERVILLE STUART ;KNEE MICHAEL JAMES (GB); SNELL & WILCOX LTD (GB) 23 octobre 1997 (1997-10-23) page 3, ligne 27 -page 5, ligne 7	2,28
A	NEWGLOWSKI J ET AL: "MOTION COMPENSATED VIDEO SEQUENCE INTERPOLATION USING DIGITAL IMAGEWARPING" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP). I. IMAGE AND MULTIDIMENSIONAL SIGNAL PROCESSING. ADELAIDE, APR. 19 - 22, 1994, NEW YORK, IEEE, US, vol. 5 CONF. 19, 19 avril 1994 (1994-04-19), pages V-205-V-208, XP000522474 ISBN: 0-7803-1776-9 alinéas 'OIII!', 'VI.A! -/--	1-37

RAPPORT DE RECHERCHE INTERNATIONALE

 De n° internationale No
 PCT/FR 02/00322

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	LECHAT P ET AL: "COMBINED MESH BASED IMAGE REPRESENTATION AND MOTION ESTIMATION, APPLICATION TO VIDEO CODING" PROCEEDINGS OF THE 1998 INTERNATIONAL CONFERENCE ON IMAGE PROCESSING. ICIP '98. CHICAGO, IL, OCT. 4 - 7, 1998, INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, LOS ALAMITOS, CA: IEEE COMPUTER SOC, US, vol. 2 CONF. 5, 4 octobre 1998 (1998-10-04), pages 909-913, XP000827604 ISBN: 0-8186-8822-X alinéas '0002!', '0004!', '0005!; figure 5	1-37
A	US 4 383 272 A (NETRAVALI ARUN N ET AL) 10 mai 1983 (1983-05-10) colonne 3, ligne 24 -colonne 4, ligne 13; figure 2	1-37
A	DEPOMMIER R ET AL: "MOTION ESTIMATION WITH DETECTION OF OCCLUSION AREAS" MULTIDIMENSIONAL SIGNAL PROCESSING. SAN FRANCISCO, MAR. 23 - 26, 1992, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP), NEW YORK, IEEE, US, vol. 3 CONF. 17, 23 mars 1992 (1992-03-23), pages 269-272, XP000378925 ISBN: 0-7803-0532-9 alinéa '0001!; figure 1	1-37
A	LAURENT N: "Hierarchical mesh-based global motion estimation, including occlusion areas detection" INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP 2000);VANCOUVER, BC, CANADA SEP 10-13 2000, vol. 3, 2000, pages 620-623, XP002201979 IEEE Int Conf Image Process;IEEE International Conference on Image Processing 2000 le document en entier	1-37
A	GB 2 279 531 A (SONY UK LTD) 4 janvier 1995 (1995-01-04) page 9, ligne 30 -page 10, ligne 20; figure 6	1-37

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

De n° de l'entrée internationale No

PCT/FR 02/00322

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
EP 0771117 A	02-05-1997	KR 209793 B1	15-07-1999
		CN 1152844 A	25-06-1997
		EP 0771117 A2	02-05-1997
		JP 9191460 A	22-07-1997
		US 5668600 A	16-09-1997
US 5978030 A	02-11-1999	KR 181034 B1	01-05-1999
		AT 207279 T	15-11-2001
		AU 711311 B2	07-10-1999
		AU 2420695 A	08-10-1996
		BR 9510541 A	09-06-1998
		CA 2215386 A1	26-09-1996
		CZ 9702902 A3	18-02-1998
		DE 69523340 D1	22-11-2001
		DE 69523340 T2	02-05-2002
		EP 0815690 A1	07-01-1998
		FI 973715 A	17-11-1997
		JP 11506576 T	08-06-1999
		WO 9629828 A1	26-09-1996
		NO 974209 A	31-10-1997
		PL 322302 A1	19-01-1998
WO 9739577 A	23-10-1997	AU 2518497 A	07-11-1997
		EP 0893025 A1	27-01-1999
		WO 9739577 A1	23-10-1997
US 4383272 A	10-05-1983	AUCUN	
GB 2279531 A	04-01-1995	JP 8009339 A	12-01-1996

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(72)発明者 ローラン - シャトネ, ナタリー

フランス国, 3 5 6 3 0 ヴィニョク, リュ・デ・フレーシュ 3

(72)発明者 ビュイソン, アレクサンドル

フランス国, 3 5 0 0 0 レンヌ, リュ・ノエル・ブラヨ 1 8, レジダンス「オンディーヌ」,
アパルトマン ベ 2 1 1